Install a stretch MetroCluster configuration

ONTAP MetroCluster

NetApp
February 13, 2026

This PDF was generated from https://docs.netapp.com/us-en/ontap-metrocluster/install-stretch/index.html on February 13, 2026. Always check docs.netapp.com for the latest.

Table of Contents

Install a stretch MetroCluster configuration

Overview

To install your stretch MetroCluster configuration, you must perform a number of procedures in the correct order.

- Prepare for the installation and understand all requirements
- Choose the correct installation procedure
- Cable the components
  - Two-node SAS-attached configuration
  - Two-node bridge-attached configuration
- Configure the software
- Test the configuration

Prepare for the MetroCluster installation

Differences between the ONTAP MetroCluster configurations

The various MetroCluster configurations have key differences in the required components.

In all configurations, each of the two MetroCluster sites are configured as an ONTAP cluster. In a two-node MetroCluster configuration, each node is configured as a single-node cluster.

Feature	**IP configurations**	**Fabric attached configurations**		**Stretch configurations**	
		Four- or eight-node	**Two-node**	**Two-node bridge-attached**	**Two-node direct-attached**
Number of controllers	Four or eight[1]	Four or eight	Two	Two	Two
Uses an FC switch storage fabric	No	Yes	Yes	No	No
Uses an IP switch storage fabric	Yes	No	No	No	No
Uses FC-to-SAS bridges	No	Yes	Yes	Yes	No

Uses direct-attached SAS storage	Yes (local attached only)	No	No	No	Yes
Supports ADP	Yes (beginning with ONTAP 9.4)	No	No	No	No
Supports local HA	Yes	Yes	No	No	No
Supports ONTAP automatic unplanned switchover (AUSO)	No	Yes	Yes	Yes	Yes
Supports unmirrored aggregates	Yes (beginning with ONTAP 9.8)	Yes	Yes	Yes	Yes
Supports ONTAP Mediator	Yes (beginning with ONTAP 9.7)	No	No	No	No
Supports MetroCluster Tiebreaker	Yes (not in combination with ONTAP Mediator)	Yes	Yes	Yes	Yes
Supports All SAN Arrays	Yes	Yes	Yes	Yes	Yes

Notes

1. Review the following considerations for eight-node MetroCluster IP configurations:
   - Eight-node configurations are supported beginning with ONTAP 9.9.1.
   - Only NetApp-validated MetroCluster switches (ordered from NetApp) are supported.
   - Configurations using IP-routed (layer 3) backend connections are not supported.

Support for All SAN Array systems in MetroCluster configurations

Some of the All SAN Arrays (ASAs) are supported in MetroCluster configurations. In the MetroCluster documentation, the information for AFF models applies to the corresponding ASA system. For example, all cabling and other information for the AFF A400 system also applies to the ASA AFF A400 system.

Supported platform configurations are listed in the NetApp Hardware Universe.

Cluster peering

Each MetroCluster site is configured as a peer to its partner site. You must be familiar with the prerequisites and guidelines for configuring the peering relationships. This is important when deciding on whether to use shared or dedicated ports for those relationships.

Related information

Cluster and SVM peering express configuration

Prerequisites for cluster peering

Before you set up cluster peering, you should confirm that connectivity between port, IP address, subnet, firewall, and cluster-naming requirements are met.

Connectivity requirements

Every intercluster LIF on the local cluster must be able to communicate with every intercluster LIF on the remote cluster.

Although it is not required, it is typically simpler to configure the IP addresses used for intercluster LIFs in the same subnet. The IP addresses can reside in the same subnet as data LIFs, or in a different subnet. The subnet used in each cluster must meet the following requirements:

- The subnet must have enough IP addresses available to allocate to one intercluster LIF per node.

  For example, in a four-node cluster, the subnet used for intercluster communication must have four available IP addresses.

Each node must have an intercluster LIF with an IP address on the intercluster network.

Intercluster LIFs can have an IPv4 address or an IPv6 address.

> ⓘ ONTAP 9 enables you to migrate your peering networks from IPv4 to IPv6 by optionally allowing both protocols to be present simultaneously on the intercluster LIFs. In earlier releases, all intercluster relationships for an entire cluster were either IPv4 or IPv6. This meant that changing protocols was a potentially disruptive event.

Port requirements

You can use dedicated ports for intercluster communication, or share ports used by the data network. Ports must meet the following requirements:

- All ports used to communicate with a given remote cluster must be in the same IPspace.

  You can use multiple IPspaces to peer with multiple clusters. Pair-wise full-mesh connectivity is required only within an IPspace.

- The broadcast domain used for intercluster communication must include at least two ports per node so that intercluster communication can fail over from one port to another port.

  Ports added to a broadcast domain can be physical network ports, VLANs, or interface groups (ifgrps).

- All ports must be cabled.
- All ports must be in a healthy state.
- The MTU settings of the ports must be consistent.

Firewall requirements

Firewalls and the intercluster firewall policy must allow the following protocols:

- ICMP service
- TCP to the IP addresses of all the intercluster LIFs over the ports 10000, 11104, and 11105
- Bidirectional HTTPS between the intercluster LIFs

The default intercluster firewall policy allows access through the HTTPS protocol and from all IP addresses (0.0.0.0/0). You can modify or replace the policy if necessary.

Considerations when using dedicated ports

When determining whether using a dedicated port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether using a dedicated port is the best intercluster network solution:

- If the amount of available WAN bandwidth is similar to that of the LAN ports, and the replication interval is such that replication occurs while regular client activity exists, then you should dedicate Ethernet ports for intercluster replication to avoid contention between replication and the data protocols.
- If the network utilization generated by the data protocols (CIFS, NFS, and iSCSI) is such that the network utilization is above 50 percent, then dedicate ports for replication to allow for nondegraded performance if a node failover occurs.
- When physical 10 GbE or faster ports are used for data and replication, you can create VLAN ports for replication and dedicate the logical ports for intercluster replication.

  The bandwidth of the port is shared between all VLANs and the base port.
- Consider the data change rate and replication interval and whether the amount of data, that must be replicated on each interval, requires enough bandwidth. This might cause contention with data protocols if sharing data ports.

Considerations when sharing data ports

When determining whether sharing a data port for intercluster replication is the correct intercluster network solution, you should consider configurations and requirements such as LAN type, available WAN bandwidth, replication interval, change rate, and number of ports.

Consider the following aspects of your network to determine whether sharing data ports is the best intercluster connectivity solution:

- For a high-speed network, such as a 40-Gigabit Ethernet (40-GbE) network, a sufficient amount of local LAN bandwidth might be available to perform replication on the same 40-GbE ports that are used for data access.

In many cases, the available WAN bandwidth is far less than the 10 GbE LAN bandwidth.

- All nodes in the cluster might have to replicate data and share the available WAN bandwidth, making data port sharing more acceptable.
- Sharing ports for data and replication eliminates the extra port counts required to dedicate ports for replication.
- The maximum transmission unit (MTU) size of the replication network will be the same size as that used on the data network.
- Consider the data change rate and replication interval and whether the amount of data, that must be replicated on each interval, requires enough bandwidth. This might cause contention with data protocols if sharing data ports.
- When data ports for intercluster replication are shared, the intercluster LIFs can be migrated to any other intercluster-capable port on the same node to control the specific data port that is used for replication.

Considerations when using unmirrored aggregates

Considerations when using unmirrored aggregates

If your configuration includes unmirrored aggregates, you must be aware of potential access issues that follow switchover operations.

Considerations for unmirrored aggregates when doing maintenance requiring power shutdown

If you are performing a negotiated switchover for maintenance reasons requiring site-wide power shutdown, you should first manually take offline any unmirrored aggregates owned by the disaster site.

If you do not take any unmirrored aggregates offline, nodes at the surviving site might go down due to multi-disk panics. This could occur if switched over unmirrored aggregates go offline, or are missing, because of the loss of connectivity to storage at the disaster site. This is the result of a power shutdown or a loss of ISLs.

Considerations for unmirrored aggregates and hierarchical namespaces

If you are using hierarchical namespaces, you should configure the junction path so that all of the volumes in that path are either on mirrored aggregates only or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates in the junction path might prevent access to the unmirrored aggregates after the switchover operation.

Considerations for unmirrored aggregates and CRS metadata volume and data SVM root volumes

The configuration replication service (CRS) metadata volume and data SVM root volumes must be on a mirrored aggregate. You cannot move these volumes to an unmirrored aggregate. If they are on an unmirrored aggregate, negotiated switchover and switchback operations are vetoed. The MetroCluster check command provides a warning if this is the case.

Considerations for unmirrored aggregates and SVMs

SVMs should be configured on mirrored aggregates only, or on unmirrored aggregates only. Configuring a mix of unmirrored and mirrored aggregates can result in a switchover operation that exceeds 120 seconds and result in a data outage if the unmirrored aggregates do not come online.

Considerations for unmirrored aggregates and SAN

In ONTAP versions prior to 9.9.1, a LUN should not be located on an unmirrored aggregate. Configuring a LUN on an unmirrored aggregate can result in a switchover operation that exceeds 120 seconds and a data outage.

Firewall usage at MetroCluster sites

Considerations for firewall usage at MetroCluster sites

If you are using a firewall at a MetroCluster site, you must ensure access for required ports.

The following table shows TCP/UDP port usage in an external firewall positioned between two MetroCluster sites.

Traffic type	Port/services
Cluster peering	11104 / TCP 11105 / TCP
ONTAP System Manager	443 / TCP
MetroCluster IP intercluster LIFs	65200 / TCP 10006 / TCP and UDP
Hardware assist	4444 / TCP

Choosing the correct installation procedure for your configuration

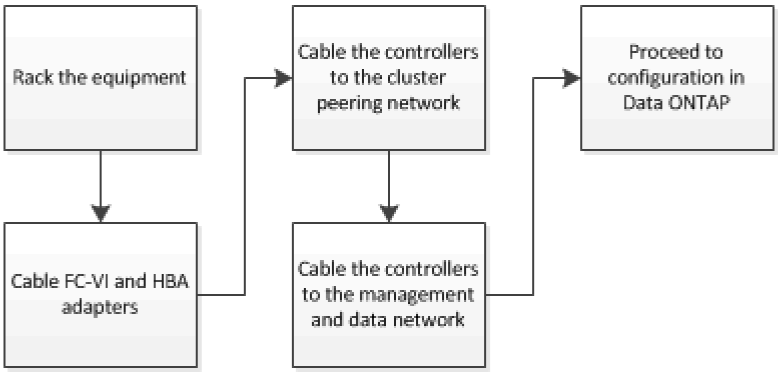
You must choose the correct installation procedure based on how the storage controllers connect to storage shelves.

For this installation type…	Use these procedures…
Two-node stretch configuration with FC-to SAS bridges	1. Cabling a two-node bridge-attached stretch MetroCluster configuration 2. Configuring the MetroCluster software in ONTAP
Two-node stretch configuration with direct-attached SAS cabling	1. Cabling a two-node SAS-attached stretch MetroCluster configuration 2. Configuring the MetroCluster software in ONTAP

Cable a two-node SAS-attached stretch MetroCluster configuration

Cabling a two-node SAS-attached stretch MetroCluster configuration

The MetroCluster components must be physically installed, cabled, and configured at both geographic sites.



Parts of a two-node SAS-attached stretch MetroCluster configuration

The two-node MetroCluster SAS-attached configuration requires a number of parts, including two single-node clusters in which the storage controllers are directly connected to the storage using SAS cables.

The MetroCluster configuration includes the following key hardware elements:

- Storage controllers

  The storage controllers connect directly to the storage using SAS cables.

  Each storage controller is configured as a DR partner to a storage controller on the partner site.

  - Copper SAS cables can be used for shorter distances.
  - Optical SAS cables can be used for longer distances.

    NetApp Interoperability Matrix Tool

    In the IMT, you can use the Storage Solution field to select your MetroCluster solution. You use the **Component Explorer** to select the components and ONTAP version to refine your search. You can click **Show Results** to display the list of supported configurations that match the criteria.

- Cluster peering network

  The cluster peering network provides connectivity for mirroring of the storage virtual machine (SVM) configuration. The configuration of all SVMs on one cluster is mirrored to the partner cluster.

Required MetroCluster hardware components and naming guidelines for two-node SAS-attached stretch configurations

The MetroCluster configuration requires a variety of hardware components. For convenience and clarity, standard names for components are used throughout the MetroCluster documentation. One site is referred to as Site A and the other site is referred to as Site B.

Supported software and hardware

The hardware and software must be supported for the MetroCluster FC configuration.

NetApp Hardware Universe

When using AFF systems, all controller modules in the MetroCluster configuration must be configured as AFF systems.

Hardware redundancy in the MetroCluster configuration

Because of the hardware redundancy in the MetroCluster configuration, there are two of each components at each site. The sites are arbitrarily assigned the letters A and B and the individual components are arbitrarily assigned the numbers 1 and 2.

Two single-node ONTAP clusters

The SAS-attached stretch MetroCluster configuration requires two single-node ONTAP clusters.

Naming must be unique within the MetroCluster configuration.

Example names:

- Site A: cluster_A
- Site B: cluster_B

Two storage controller modules

The SAS-attached stretch MetroCluster configuration requires two storage controller modules.

- Naming must be unique within the MetroCluster configuration.
- All controller modules in the MetroCluster configuration must be running the same version of ONTAP.
- All controller modules in a DR group must be of the same model.
- All controller modules in a DR group must use the same FC-VI configuration.

  Some controller modules support two options for FC-VI connectivity:

  - Onboard FC-VI ports
  - An FC-VI card in slot 1

    A mix of one controller module using onboard FC-VI ports and another using an add-on FC-VI card is not supported. For example, if one node uses onboard FC-VI configuration, then all other nodes in the DR group must use onboard FC-VI configuration as well.

Example names:

- Site A: controller_A_1
- Site B: controller_B_1

At least four SAS disk shelves (recommended)

The SAS-attached stretch MetroCluster configuration requires at least two SAS disk shelves. Four SAS disk shelves is recommended.

Two shelves are recommended at each site to allow disk ownership on a per-shelf basis. A minimum of one shelf at each site is supported.

Example names:

- Site A:
  - shelf_A_1_1
  - shelf_A_1_2
- Site B:
  - shelf_B_1_1
  - shelf_B_1_2

Install and cable MetroCluster components for two-node SAS-attached stretch configurations

Installing and cabling MetroCluster components for two-node SAS-attached stretch configurations

The storage controllers must be cabled to the storage media and to each other. The storage controllers must also be cabled to the data and management network.

Before you begin any procedure in this document

The following overall requirements must be met before completing this task:

- Prior to installation you must have familiarized yourself with the considerations and best practices for installing and cabling disk shelves for your disk shelf model.
- All MetroCluster components must be supported.

  NetApp Interoperability Matrix Tool

  In the IMT, you can use the Storage Solution field to select your MetroCluster solution. Use the **Component Explorer** to select the components and ONTAP version to refine your search. You can click **Show Results** to display the list of supported configurations that match the criteria.

About this task

- The terms node and controller are used interchangeably.

Racking the hardware components

If you have not received the equipment already installed in cabinets, you must rack the components.

This task must be performed on both MetroCluster sites.

Steps

1. Plan the positioning of the MetroCluster components.

   The amount of rack space needed depends on the platform model of the storage controllers, the switch types, and the number of disk shelf stacks in your configuration.

2. Using standard shop practices for working with electrical equipment make sure you are properly grounded.
3. Install the storage controllers in the rack or cabinet.

   ONTAP Hardware Systems Documentation

4. Install the disk shelves, daisy-chain the disk shelves in each stack, power them on, and set the shelf IDs.

   See the appropriate guide for your disk shelf model for information about daisy-chaining disk shelves and setting shelf IDs.

   > Shelf IDs must be unique for each SAS disk shelf within each MetroCluster DR group (including both sites). When manually setting shelf IDs, you must power-cycle the disk shelf.

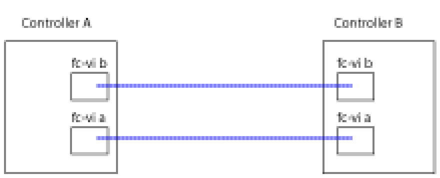
Cabling the controllers to each other and the storage shelves

The controller FC-VI adapters must be cabled directly to each other. The controller SAS ports must be cabled to both the remote and local storage stacks.

This task must be performed at both MetroCluster sites.

Steps

1. Cable the FC-VI ports.

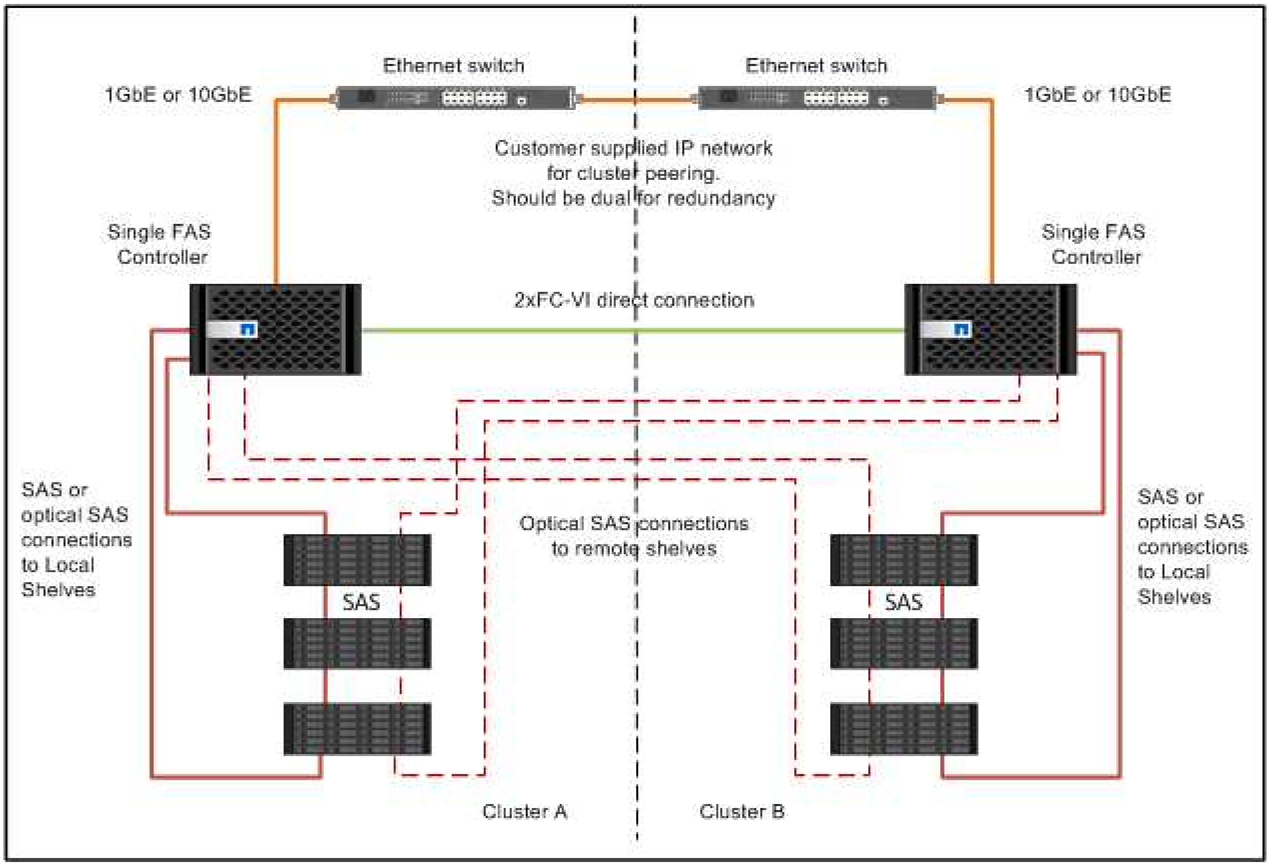
   

   The above illustration is a typical representative cable connection. The specific FC-VI ports will vary by controller module.

   - FAS8200 and AFF A300 controller modules can be ordered with one of two options for FC-VI connectivity:
     - Onboard ports 0e and 0f are configured in FC-VI mode.
     - Ports 1a and 1b on an FC-VI card go in slot 1.
   - AFF A700 and FAS9000 storage systems controller modules use four FC-VI ports each.
   - AFF A400 and FAS8300 storage system controller modules use FC-VI ports 2a and 2b.

2. Cable the SAS ports.

   The following illustration shows the connections. Your port usage might be different depending on the available SAS and FC-VI ports on the controller module.



Cabling the cluster peering connections

You must cable the controller module ports used for cluster peering so that they have connectivity with the cluster on their partner site.

This task must be performed on each controller module in the MetroCluster configuration.

At least two ports on each controller module should be used for cluster peering.

The recommended minimum bandwidth for the ports and network connectivity is 1 GbE.

Steps

1. Identify and cable at least two ports for cluster peering and verify they have network connectivity with the partner cluster.

   Cluster peering can be done on dedicated ports or on data ports. Using dedicated ports provides a higher throughput for the cluster peering traffic.

   Cluster and SVM peering express configuration

Cabling the management and data connections

You must cable the management and data ports on each storage controller to the site networks.

This task must be repeated for each new controller at both MetroCluster sites.

You can connect the controller and cluster switch management ports to existing switches in your network. In addition you can connect controller to new dedicated network switches such as NetApp CN1601 cluster management switches.

Steps

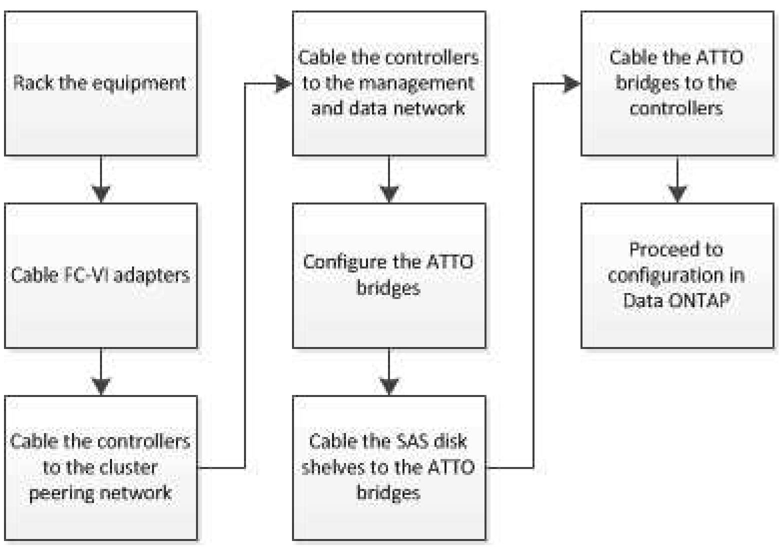
1. Cable the controller's management and data ports to the management and data networks at the local site.

   ONTAP Hardware Systems Documentation

Cable a two-node bridge-attached stretch MetroCluster configuration

Cabling a two-node bridge-attached stretch MetroCluster configuration

The MetroCluster components must be physically installed, cabled, and configured at both geographic sites.



Parts of a two-node bridge-attached stretch MetroCluster configuration

As you plan your MetroCluster configuration, you should understand the parts of the configuration and how they work together.

The MetroCluster configuration includes the following key hardware elements:

- Storage controllers

  The storage controllers are not connected directly to the storage but connected to FC-to-SAS bridges. The storage controllers are connected to each other by FC cables between each controller's FC-VI adapters.

Each storage controller is configured as a DR partner to a storage controller on the partner site.

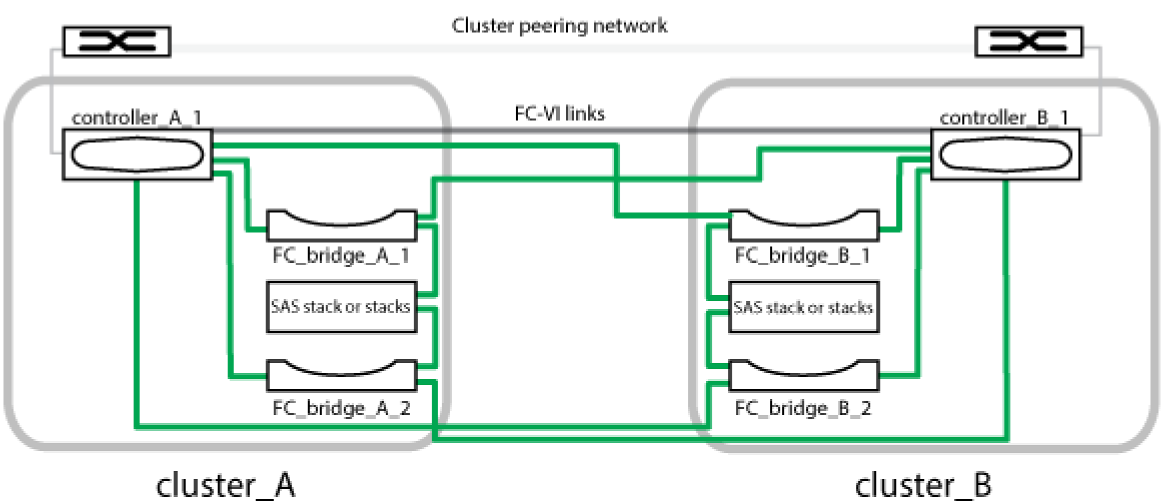
- FC-to-SAS bridges

  The FC-to-SAS bridges connect the SAS storage stacks to the FC initiator ports on the controllers, providing bridging between the two protocols.

- Cluster peering network

  The cluster peering network provides connectivity for mirroring of the storage virtual machine (SVM) configuration. The configuration of all SVMs on one cluster is mirrored to the partner cluster.

The following illustration shows a simplified view of the MetroCluster configuration. For some connections, a single line represents multiple, redundant connections between the components. Data and management network connections are not shown.



- The configuration consists of two single-node clusters.
- Each site has one or more stacks of SAS storage.

  SAS shelves in MetroCluster configurations are not supported with ACP cabling.

  Additional storage stacks are supported, but only one is shown at each site.

Required MetroCluster hardware components and naming conventions for two-node bridge-attached stretch configurations

When planning your MetroCluster configuration, you must understand the required and supported hardware and software components. For convenience and clarity, you should also understand the naming conventions used for components in examples throughout the documentation. For example, one site is referred to as Site A and the other site is referred to as Site B.

Supported software and hardware

The hardware and software must be supported for the MetroCluster FC configuration.

NetApp Hardware Universe

When using AFF systems, all controller modules in the MetroCluster configuration must be configured as AFF systems.

Hardware redundancy in the MetroCluster configuration

Because of the hardware redundancy in the MetroCluster configuration, there are two of each component at each site. The sites are arbitrarily assigned the letters A and B and the individual components are arbitrarily assigned the numbers 1 and 2.

Requirement for two single-node ONTAP clusters

The bridge-attached stretch MetroCluster configuration requires two single-node ONTAP clusters.

Naming must be unique within the MetroCluster configuration.

Example names:

- Site A: cluster_A
- Site B: cluster_B

Requirement for two storage controller modules

The bridge-attached stretch MetroCluster configuration requires two storage controller modules.

The controllers must meet the following requirements:

- Naming must be unique within the MetroCluster configuration.
- All controller modules in the MetroCluster configuration must be running the same version of ONTAP.
- All controller modules in a DR group must be of the same model.
- All controller modules in a DR group must use the same FC-VI configuration.

  Some controller modules support two options for FC-VI connectivity:

  - Onboard FC-VI ports
  - An FC-VI card in slot 1

    A mix of one controller module using onboard FC-VI ports and another using an add-on FC-VI card is not supported. For example, if one node uses onboard FC-VI configuration, then all other nodes in the DR group must use onboard FC-VI configuration as well.

Example names:

- Site A: controller_A_1
- Site B: controller_B_1

Requirement for FC-to-SAS bridges

The bridge-attached stretch MetroCluster configuration requires two or more FC-to-SAS bridges at each site.

These bridges connect the SAS disk shelves to the controller modules.

FibreBridge 6500N bridges are not supported in configurations running ONTAP 9.8 and later.

- FibreBridge 7600N and 7500N bridges support up to four SAS stacks.
- Each stack can use different models of IOM, but all shelves within a stack must use the same model.

  The supported IOM models depend on the ONTAP version you are running.
- Naming must be unique within the MetroCluster configuration.

The suggested names used as examples in this procedure identify the controller module that the bridge connects to and the port.

Example names:

- Site A:
  - bridge_A_1_*port-number*
  - bridge_A_2_*port-number*
- Site B:
  - bridge_B_1_*port-number*
  - bridge_B_2_*port-number*

Requirement for at least four SAS shelves (recommended)

The bridge-attached stretch MetroCluster configuration requires at least two SAS shelves. However, two shelves are recommended at each site to allow disk ownership on a per-shelf basis, for a total of four SAS shelves.

A minimum of one shelf at each site is supported.

Example names:

- Site A:
  - shelf_A_1_1
  - shelf_A_1_2
- Site B:
  - shelf_B_1_1
  - shelf_B_1_2

Information gathering worksheet for FC-to-SAS bridges

Before beginning to configure the MetroCluster sites, you should gather required configuration information.

Site A, FC-to-SAS bridge 1 (FC_bridge_A_1a)

Each SAS stack requires at least two FC-to-SAS bridges.

Each bridge connects to Controller_A_1_*port-number* and Controller_B_1_*port-number*.

Site A	Your value
Bridge_A_1a IP address	
Bridge_A_1a Username	
Bridge_A_1a Password	

Site A, FC-to-SAS bridge 2 (FC_bridge_A_1b)

Each SAS stack requires at least two FC-to-SAS bridges.

Each bridge connects to Controller_A_1_`port-number` and Controller_B_1_`port-number`.

Site A	Your value
Bridge_A_1b IP address	
Bridge_A_1b Username	
Bridge_A_1b Password	

Site B, FC-to-SAS bridge 1 (FC_bridge_B_1a)

Each SAS stack requires at least two FC-to-SAS bridges.

Each bridge connects to Controller_A_1_`*port-number*` and Controller_B_1_`*port-number*`.

Site B	Your value
Bridge_B_1a IP address	
Bridge_B_1a Username	
Bridge_B_1a Password	

Site B, FC-to-SAS bridge 2 (FC_bridge_B_1b)

Each SAS stack requires at least two FC-to-SAS bridges.

Each bridge connects to Controller_A_1_`*port-number*` and Controller_B_1_`*port-number*`.

Site B	Your value
Bridge_B_1b IP address	
Bridge_B_1b Username	

Bridge_B_1b Password	

Install and cable MetroCluster components

Racking the hardware components

If you have not received the equipment already installed in cabinets, you must rack the components.

This task must be performed on both MetroCluster sites.

Steps

1. Plan out the positioning of the MetroCluster components.

   The rack space depends on the platform model of the storage controllers, switch types, and the number of disk shelf stacks in your configuration.

2. Properly ground yourself.
3. Install the storage controllers in the rack or cabinet.

   ONTAP Hardware Systems Documentation

4. Install the disk shelves, power them on, and set the shelf IDs.
   - You must power-cycle each disk shelf.
   - Shelf IDs must be unique for each SAS disk shelf within each MetroCluster DR group (including both sites).
5. Install each FC-to-SAS bridge:
   a. Secure the “L” brackets on the front of the bridge to the front of the rack (flush-mount) with the four screws.

      The openings in the bridge “L” brackets are compliant with rack standard ETA-310-X for 19-inch (482.6 mm) racks.

      For more information and an illustration of the installation, see the *ATTO FibreBridge Installation and Operation Manual for your bridge model*.

   b. Connect each bridge to a power source that provides a proper ground.
   c. Power on each bridge.

      > For maximum resiliency, bridges that are attached to the same stack of disk shelves must be connected to different power sources.

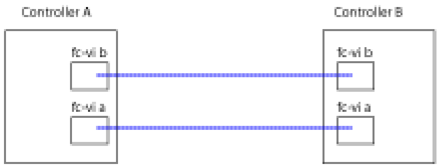
      The bridge Ready LED might take up to 30 seconds to illuminate, indicating that the bridge has completed its power-on self test sequence.

Cabling the controllers to each other

Each controller’s FC-VI adapters must be cabled directly to its partner.

Steps

1. Cable the FC-VI ports.



The above illustration is a typical representation of the required cabling. The specific FC-VI ports vary by controller module.

   - AFF A300 and FAS8200 controller modules can be ordered with one of two options for FC-VI connectivity:
     - Onboard ports 0e and 0f configured in FC-VI mode.
     - Ports 1a and 1b on an FC-VI card in slot 1.
   - AFF A700 and FAS9000 storage systems controller modules use four FC-VI ports each.

Cabling the cluster peering connections

You must cable the controller module ports used for cluster peering so that they have connectivity with the cluster on their partner site.

This task must be performed on each controller module in the MetroCluster configuration.

At least two ports on each controller module should be used for cluster peering.

The recommended minimum bandwidth for the ports and network connectivity is 1 GbE.

Steps

1. Identify and cable at least two ports for cluster peering and verify they have network connectivity with the partner cluster.

   Cluster peering can be done on dedicated ports or on data ports. Using dedicated ports provides a higher throughput for the cluster peering traffic.

   Cluster and SVM peering express configuration

Cabling the management and data connections

You must cable the management and data ports on each storage controller to the site networks.

This task must be repeated for each new controller at both MetroCluster sites.

You can connect the controller and cluster switch management ports to existing switches in your network. In addition you can connect controller to new dedicated network switches such as NetApp CN1601 cluster management switches.

Steps

1. Cable the controller’s management and data ports to the management and data networks at the local site.

Install FC-to-SAS bridges and SAS disk shelves

Install and cable ATTO FibreBridge bridges and SAS disk shelves when you add new storage to the configuration.

About this task

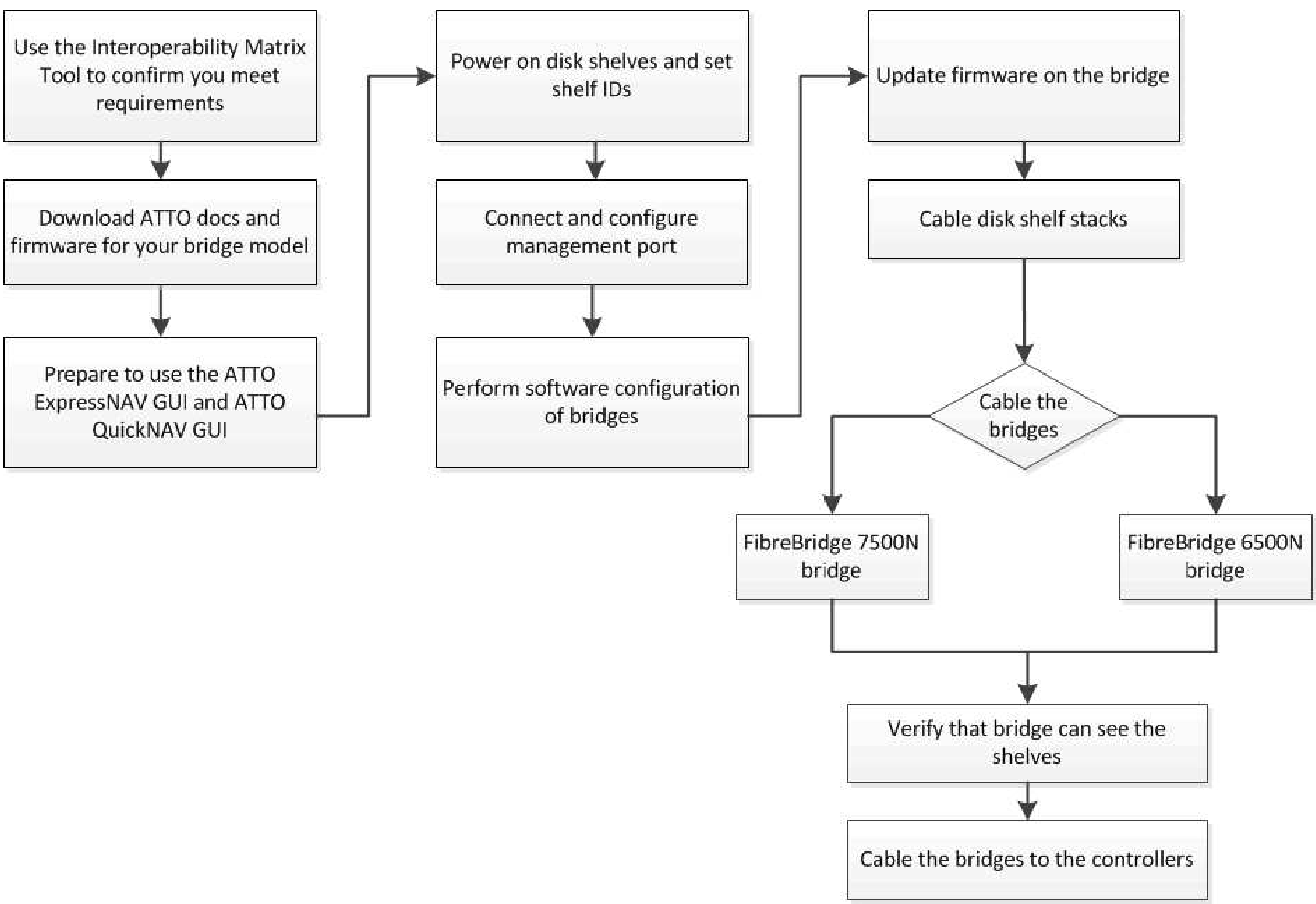
For systems received from the factory, the FC-to-SAS bridges are preconfigured and do not require additional configuration.

This procedure is written with the assumption that you are using the recommended bridge management interfaces: the ATTO ExpressNAV GUI and ATTO QuickNAV utility.

You use the ATTO ExpressNAV GUI to configure and manage a bridge, and to update the bridge firmware. You use the ATTO QuickNAV utility to configure the bridge Ethernet management 1 port.

You can use other management interfaces instead, if needed, such as a serial port or Telnet to configure and manage a bridge and to configure the Ethernet management 1 port, and FTP to update the bridge firmware.

This procedure uses the following workflow:



In-band management of the FC-to-SAS bridges

Beginning with ONTAP 9.5 with FibreBridge 7500N or 7600N bridges, *in-band management* of the bridges is supported as an alternative to IP management of the bridges. Beginning with ONTAP 9.8, out-of-band management is deprecated.

> Beginning with ONTAP 9.8, the `storage bridge` command is replaced with `system bridge`. The following steps show the `storage bridge` command, but if you are running ONTAP 9.8 or later, the `system bridge` command is preferred.

When using in-band management, the bridges can be managed and monitored from the ONTAP CLI using the FC connection to the bridge. Physical access to the bridge through the bridge Ethernet ports is not required, reducing the security vulnerability of the bridge.

The availability of in-band management of the bridges depends on the version of ONTAP:

- Beginning with ONTAP 9.8, bridges are managed via in-band connections by default and out-of-band management of the bridges via SNMP is deprecated.
- ONTAP 9.5 through 9.7: Either in-band management or out-of-band SNMP management is supported.
- Before ONTAP 9.5, only out-of-band SNMP management is supported.

Bridge CLI commands can be issued from the ONTAP interface `storage bridge run-cli -name <bridge_name> -command <bridge_command_name>` command at the ONTAP interface.

> Using in-band management with IP access disabled is recommended to improve security by limiting physical connectivity the bridge.

FibreBridge 7600N and 7500N bridge limits and attachment rules

Review the limits and considerations when attaching FibreBridge 7600N and 7500N bridges.

FibreBridge 7600N and 7500N bridge limits

- The maximum number of HDD and SSD drives combined is 240.
- The maximum number of SSD drives is 96.
- The maximum number of SSDs per SAS port is 48.
- The maximum number of shelves per SAS port is 10.

FibreBridge 7600N and 7500N bridge attachment rules

- Do not mix SSD and HDD drives on the same SAS port.
- Distribute the shelves evenly across the SAS ports.
- You shouldn’t have DS460 shelves on the same SAS port as other shelf types (for example, DS212 or DS224 shelves).

Example configuration

The following shows an example configuration for connecting four DS224 shelves with SSD drives and six DS224 shelves with HDD drives:

SAS port	Shelves and drives
SAS port-A	2x DS224 shelves with SSD drives
SAS port-B	2x DS224 shelves with SSD drives

SAS port	Shelves and drives
SAS port-C	3x DS224 shelves with HDD drives
SAS port-D	3x DS224 shelves with HDD drives

Prepare for the installation

When you are preparing to install the bridges as part of your new MetroCluster system, you must verify that your system meets certain requirements, including meeting setup and configuration requirements for the bridges. Other requirements include downloading the necessary documents, the ATTO QuickNAV utility, and the bridge firmware.

Before you begin

- Your system must already be installed in a rack if it was not shipped in a system cabinet.
- Your configuration must be using supported hardware models and software versions.

  In the NetApp Interoperability Matrix Tool (IMT), you can use the **Storage Solution** field to select your MetroCluster solution. You can use the **Component Explorer** to select the components and ONTAP version to refine your search. You can select **Show Results** to display the list of supported configurations that match the criteria.

- Each FC controller must have one FC port available for one bridge to connect to it.
- You must be familiar with how to handle SAS cables and the considerations and best practices for installing and cabling disk shelves.

  The *Installation and Service Guide* for your disk shelf model describes the considerations and best practices.

- The computer you are using to set up the bridges must be running an ATTO-supported web browser to use the ATTO ExpressNAV GUI.

  The *ATTO Product Release Notes* have an up-to-date list of supported web browsers. You can access this document from the ATTO web site as described in the following steps.

Steps

1. Download the *Installation and Service Guide* for your disk shelf model:
   a. Access the ATTO web site using the link provided for your FibreBridge model and download the manual and the QuickNAV utility.

      > The *ATTO FibreBridge Installation and Operation Manual* for your model bridge has more information about management interfaces.
      >
      > You can access this and other content on the ATTO web site by using the link provided on the ATTO FibreBridge Description page.

2. Gather the hardware and information needed to use the recommended bridge management interfaces, the ATTO ExpressNAV GUI, and the ATTO QuickNAV utility:
   a. Determine a non-default user name and password (for accessing the bridges).

      You should change the default user name and password.

b. If configuring for IP management of the bridges, you need the shielded Ethernet cable provided with the bridges (which connects from the bridge Ethernet management 1 port to your network).

c. If configuring for IP management of the bridges, you need an IP address, subnet mask, and gateway information for the Ethernet management 1 port on each bridge.

d. Disable VPN clients on the computer you are using for setup.

   Active VPN clients cause the QuickNAV scan for bridges to fail.

Install the FC-to-SAS bridge and SAS shelves

After ensuring that the system meets all of the requirements in “Preparing for the installation”, you can install your new system.

About this task

- The disk and shelf configuration at both sites should be identical.

  If a non-mirrored aggregate is used, the disk and shelf configuration at each site might be different.

  > All disks in the disaster recovery group must use the same type of connection and be visible to all of the nodes within the disaster recovery group, regardless of the disks being used for mirrored or non-mirrored aggregate.

- The system connectivity requirements for maximum distances for disk shelves, FC controllers, and backup tape devices using 50-micron, multimode fiber-optic cables, also apply to FibreBridge bridges.

  NetApp Hardware Universe

> In-band ACP is supported without additional cabling in the following shelves and FibreBridge 7500N or 7600N bridge:
>
> - IOM12 (DS460C) behind a 7500N or 7600N bridge with ONTAP 9.2 and later
> - IOM12 (DS212C and DS224C) behind a 7500N or 7600N bridge with ONTAP 9.1 and later

> SAS shelves in MetroCluster configurations do not support ACP cabling.

Enable IP port access on the FibreBridge 7600N bridge if necessary

If you are using an ONTAP version prior to 9.5, or otherwise plan to use out-of-band access to the FibreBridge 7600N bridge using telnet or other IP port protocols and services (FTP, ExpressNAV, ICMP, or QuickNAV), you can enable the access services via the console port.

About this task

Unlike the ATTO FibreBridge 7500N bridges, the FibreBridge 7600N bridge is shipped with all IP port protocols and services disabled.

Beginning with ONTAP 9.5, *in-band management* of the bridges is supported. This means the bridges can be configured and monitored from the ONTAP CLI via the FC connection to the bridge. Physical access to the bridge via the bridge Ethernet ports is not required and the bridge user interfaces are not required.

Beginning with ONTAP 9.8, *in-band management* of the bridges is supported by default and out-of-band SNMP management is deprecated.

This task is required if you are **not** using in-band management to manage the bridges. In this case, you need to configure the bridge via the Ethernet management port.

Steps

1. Access the bridge console interface by connecting a serial cable to the serial port on the FibreBridge 7600N bridge.
2. Using the console, enable the access services, and then save the configuration:

   `set closeport none`

   `saveconfiguration`

   The `set closeport none` command enables all access services on the bridge.

3. Disable a service, if desired, by issuing the `set closeport` command and repeating the command as necessary until all desired services are disabled:

   `set closeport` *`service`*

   The `set closeport` command disables a single service at a time.

   The parameter *`service`* can be specified as one of the following:

   - expressnav
   - ftp
   - icmp
   - quicknav
   - snmp
   - telnet

   You can check whether a specific protocol is enabled or disabled by using the `get closeport` command.

4. If you are enabling SNMP, you must also issue following command:

   `set SNMP enabled`

   SNMP is the only protocol that requires a separate enable command.

5. Save the configuration:

   `saveconfiguration`

Configure the FC-to-SAS bridges

Before cabling your model of the FC-to-SAS bridges, you must configure the settings in the FibreBridge software.

Before you begin

You should decide whether you will be using in-band management of the bridges.

Beginning with ONTAP 9.8, the `storage bridge` command is replaced with `system bridge`. The following steps show the `storage bridge` command, but if you are running ONTAP 9.8 or later, the `system bridge` command is preferred.

About this task

If you will be using in-band management of the bridge rather than IP management, the steps for configuring the Ethernet port and IP settings can be skipped, as noted in the relevant steps.

Steps

1. Configure the serial console port on the ATTO FibreBridge by setting the port speed to 115000 bauds:

```
get serialportbaudrate
SerialPortBaudRate = 115200

Ready.

set serialportbaudrate 115200

Ready. *
saveconfiguration
Restart is necessary....
Do you wish to restart (y/n) ? y
```

2. If configuring for in-band management, connect a cable from FibreBridge RS-232 serial port to the serial (COM) port on a personal computer.

   The serial connection will be used for initial configuration, and then in-band management via ONTAP and the FC ports can be used to monitor and manage the bridge.

3. If configuring for IP management, connect the Ethernet management 1 port on each bridge to your network by using an Ethernet cable.

   In systems running ONTAP 9.5 or later, in-band management can be used to access the bridge via the FC ports rather than the Ethernet port. Beginning with ONTAP 9.8, only in-band management is supported and SNMP management is deprecated.

   The Ethernet management 1 port enables you to quickly download the bridge firmware (using ATTO ExpressNAV or FTP management interfaces) and to retrieve core files and extract logs.

4. If configuring for IP management, configure the Ethernet management 1 port for each bridge by following the procedure in section 2.0 of the *ATTO FibreBridge Installation and Operation Manual* for your bridge model.

   In systems running ONTAP 9.5 or later, in-band management can be used to access the bridge via the FC ports rather than the Ethernet port. Beginning with ONTAP 9.8, only in-band management is supported and SNMP management is deprecated.

   When running QuickNAV to configure an Ethernet management port, only the Ethernet management port that is connected by the Ethernet cable is configured. For example, if you also wanted to configure the Ethernet management 2 port, you would need to connect the Ethernet cable to port 2 and run QuickNAV.

5. Configure the bridge.

   You should make note of the user name and password that you designate.

   > Do not configure time synchronization on ATTO FibreBridge 7600N or 7500N. The time synchronization for ATTO FibreBridge 7600N or 7500N is set to the cluster time after the bridge is discovered by ONTAP. It is also synchronized periodically once a day. The time zone used is GMT and is not changeable.

   a. If configuring for IP management, configure the IP settings of the bridge.

      In systems running ONTAP 9.5 or later, in-band management can be used to access the bridge via the FC ports rather than the Ethernet port. Beginning with ONTAP 9.8, only in-band management is supported and SNMP management is deprecated.

      To set the IP address without the QuickNAV utility, you need to have a serial connection to the FibreBridge.

      If using the CLI, you must run the following commands:

      ```
      set ipaddress mp1 ip-address
      ```

      ```
      set ipsubnetmask mp1 subnet-mask
      ```

      ```
      set ipgateway mp1 x.x.x.x
      ```

      ```
      set ipdhcp mp1 disabled
      ```

      ```
      set ethernetspeed mp1 1000
      ```

   b. Configure the bridge name.

      The bridges should each have a unique name within the MetroCluster configuration.

      Example bridge names for one stack group on each site:

      - bridge_A_1a
      - bridge_A_1b
      - bridge_B_1a
      - bridge_B_1b

      If using the CLI, you must run the following command:

      ```
      set bridgename <bridge_name>
      ```

   c. If running ONTAP 9.4 or earlier, enable SNMP on the bridge:

      ```
      set SNMP enabled
      ```

      In systems running ONTAP 9.5 or later, in-band management can be used to access the bridge via the FC ports rather than the Ethernet port. Beginning with ONTAP 9.8, only in-band management is supported and SNMP management is deprecated.

6. Configure the bridge FC ports.

   a. Configure the data rate/speed of the bridge FC ports.

      The supported FC data rate depends on your model bridge.

      - The FibreBridge 7600N bridge supports up to 32, 16, or 8 Gbps.
      - The FibreBridge 7500N bridge supports up to 16, 8, or 4 Gbps.

      > The FCDataRate speed you select is limited to the maximum speed supported by both the bridge and the FC port of the controller module to which the bridge port connects. Cabling distances must not exceed the limitations of the SFPs and other hardware.

      If using the CLI, you must run the following command:

      `set FCDataRate <port-number> <port-speed>`

   b. If you are configuring a FibreBridge 7500N bridge, configure the connection mode that the port uses to "ptp".

      > The FCConnMode setting is not required when configuring a FibreBridge 7600N bridge.

      If using the CLI, you must run the following command:

      `set FCConnMode <port-number> ptp`

   c. If you are configuring a FibreBridge 7600N or 7500N bridge, you must configure or disable the FC2 port.

      - If you are using the second port, you must repeat the previous substeps for the FC2 port.
      - If you are not using the second port, then you must disable the port:

        `FCPortDisable <port-number>`

        The following example shows the disabling of FC port 2:

        ```
        FCPortDisable 2

        Fibre Channel Port 2 has been disabled.
        ```

   d. If you are configuring a FibreBridge 7600N or 7500N bridge, disable the unused SAS ports:

      `SASPortDisable` *`sas-port`*

      > SAS ports A through D are enabled by default. You must disable the SAS ports that are not being used.

      If only SAS port A is used, then SAS ports B, C, and D must be disabled. The following example shows the disabling of SAS port B. You must similarly disable SAS ports C and D:

```
SASPortDisable b

SAS Port B has been disabled.
```

7. Secure access to the bridge and save the bridge’s configuration. Choose an option from below depending on the version of ONTAP your system is running.

ONTAP version	Steps

ONTAP 9.5 or later	a. View the status of the bridges:
	
	`storage bridge show`
	
	The output shows which bridge is not secured.
	
	b. Secure the bridge:
	
	`securebridge`

ONTAP 9.4 or earlier	a. View the status of the bridges:
	
	`storage bridge show`
	
	The output shows which bridge is not secured.
	
	b. Check the status of the unsecured bridge’s ports:
	
	`info`
	
	The output shows the status of Ethernet ports MP1 and MP2.
	
	c. If Ethernet port MP1 is enabled, run:
	
	`set EthernetPort mp1 disabled`
	
	If Ethernet port MP2 is also enabled, repeat the previous substep for port MP2.
	
	d. Save the bridge’s configuration.
	
	You must run the following commands:
	
	`SaveConfiguration`
	
	`FirmwareRestart`
	
	You are prompted to restart the bridge.

8. After completing MetroCluster configuration, use the `flashimages` command to check your version of FibreBridge firmware and, if the bridges are not using the latest supported version, update the firmware on

all bridges in the configuration.

Maintain MetroCluster Components

Cable a FibreBridge 7600N or 7500N bridge with disk shelves using IOM12 modules

After configuring the bridge, you can start cabling your new system.

About this task

For disk shelves, you insert a SAS cable connector with the pull tab oriented down (on the underside of the connector).

Steps

1. Daisy-chain the disk shelves in each stack:

    a. Beginning with the logical first shelf in the stack, connect IOM A port 3 to the to IOM A port 1 on the next shelf until each IOM A in the stack is connected.

    b. Repeat the previous substep for IOM B.

    c. Repeat the previous substeps for each stack.

    The *Installation and Service Guide* for your disk shelf model provides detailed information about daisy-chaining disk shelves.

2. Power on the disk shelves, and then set the shelf IDs.

    - You must power-cycle each disk shelf.
    - Shelf IDs must be unique for each SAS disk shelf within each MetroCluster DR group (including both sites).

3. Cable disk shelves to the FibreBridge bridges.

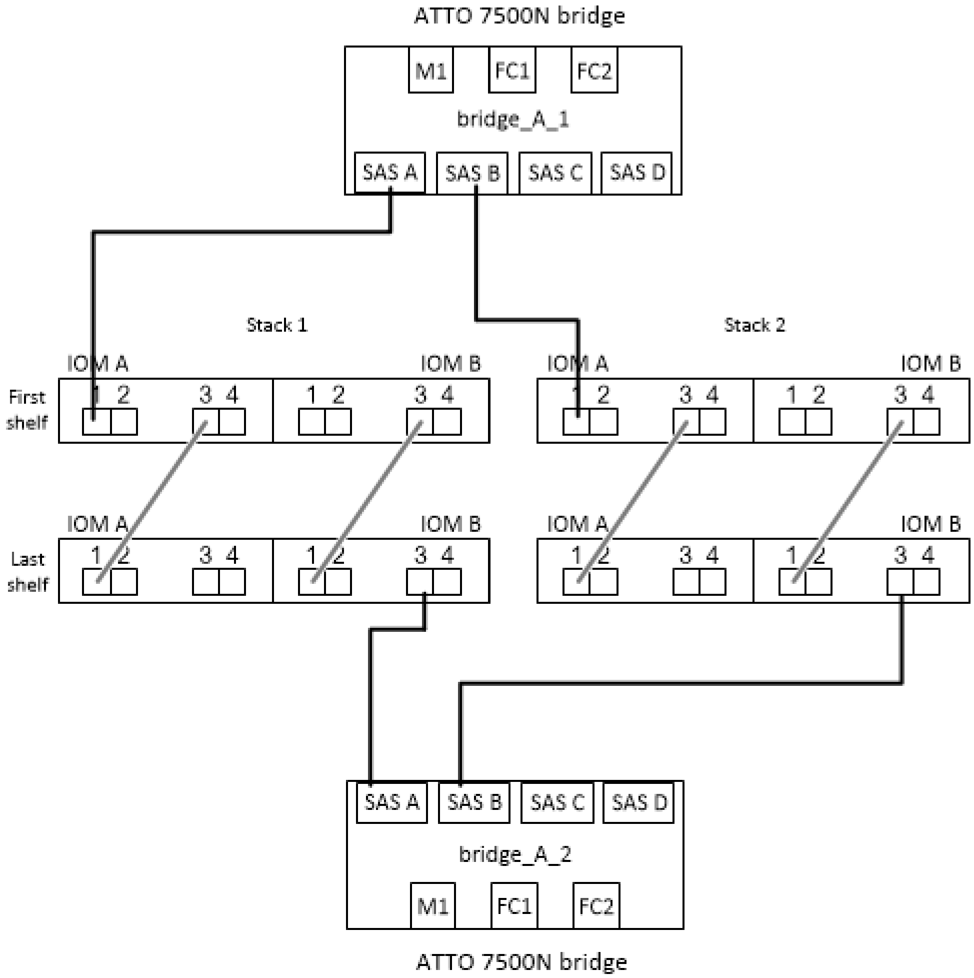
    a. For the first stack of disk shelves, cable IOM A of the first shelf to SAS port A on FibreBridge A, and cable IOM B of the last shelf to SAS port A on FibreBridge B.

    b. For additional shelf stacks, repeat the previous step using the next available SAS port on the FibreBridge bridges, using port B for the second stack, port C for the third stack, and port D for the fourth stack.

    c. During cabling, attach the stacks based on IOM12 modules to the same bridge as long as they are connected to separate SAS ports.

    > Each stack can use different models of IOM, but all disk shelves within a stack must use the same model.

    The following illustration shows disk shelves connected to a pair of FibreBridge 7600N or 7500N bridges:



Verify bridge connectivity and cable the FC-to-SAS bridges to the controller FC ports

You must cable the bridges to the controller FC ports in a two-node bridge-attached MetroCluster configuration.

Steps

1. Verify that each bridge can detect all of the disk drives and disk shelves to which the bridge is connected:

   `sastargets`

   The `sastargets` command output shows the devices (disks and disk shelves) connected to the bridge. The output lines are sequentially numbered so that you can quickly count the devices.

   The following output shows that 10 disks are connected:

```
Tgt VendorID ProductID        Type        SerialNumber
  0 NETAPP   X410_S15K6288A15 DISK        3QP1CLE300009940UHJV
  1 NETAPP   X410_S15K6288A15 DISK        3QP1ELF600009940V1BV
  2 NETAPP   X410_S15K6288A15 DISK        3QP1G3EW00009940U2M0
  3 NETAPP   X410_S15K6288A15 DISK        3QP1EWMP00009940U1X5
  4 NETAPP   X410_S15K6288A15 DISK        3QP1FZLE00009940G8YU
  5 NETAPP   X410_S15K6288A15 DISK        3QP1FZLF00009940TZKZ
  6 NETAPP   X410_S15K6288A15 DISK        3QP1CEB400009939MGXL
  7 NETAPP   X410_S15K6288A15 DISK        3QP1G7A900009939FNTT
  8 NETAPP   X410_S15K6288A15 DISK        3QP1FY0T00009940G8PA
  9 NETAPP   X410_S15K6288A15 DISK        3QP1FXW600009940VERQ
```

2. Verify that the command output shows that the bridge is connected to the correct disks and disk shelves in the stack.

If the output is…	Then…
Correct	Repeat Step 1 for each remaining bridge.
Not correct	a. Check for loose SAS cables or correct the SAS cabling by recabling the disk shelves to the bridges. Cable a FibreBridge 7600N or 7500N bridge with disk shelves using IOM12 modules b. Repeat Step 1 for each remaining bridge.

3. Cable each bridge to the controller FC ports:

   a. Cable FC port 1 of the bridge to an FC port on the controller in cluster_A.

   b. Cable FC port 2 of the bridge to an FC port on the controller in cluster_B.

      - If the controller is configured with a quad-port FC adapter, make sure that the bridges at either end of the storage stack are not connected to two FC ports on the same ASIC. For example:

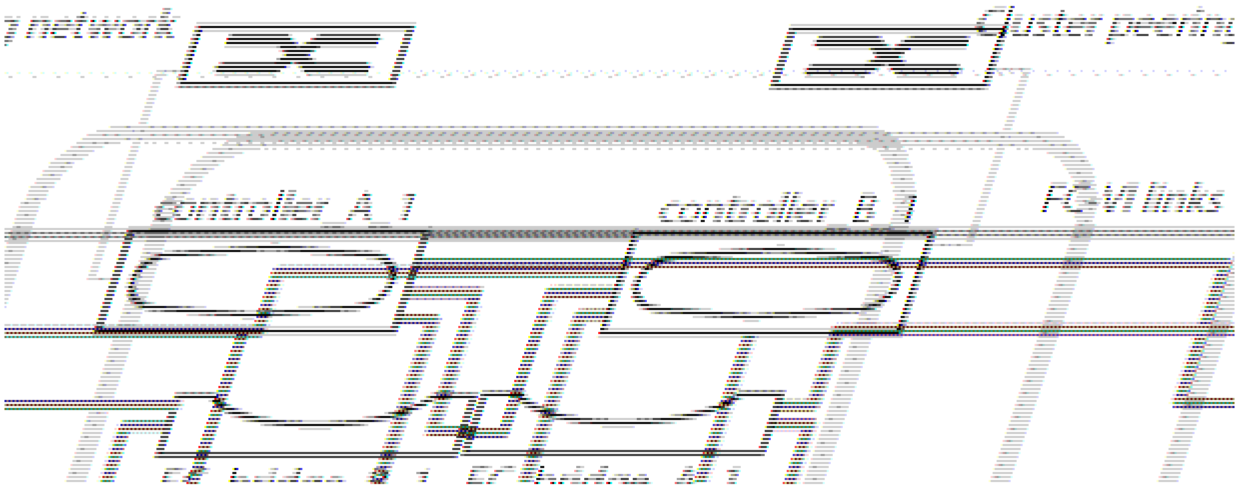
        - Port a and port b share the same ASIC.

        - Port c and port d share the same ASIC.

          In this example, connect FC_bridge_A_1 to port a and FC_bridge_A2 to port c.

      - If the controller is configured with more than one FC adapter, do not cable the bridges at either end of the storage stack to the same adapter.

        In this scenario, you should connect FC_bridge_A_1 to an onboard FC port, and connect FC_bridge_A_2 to an FC port on an adapter in an expansion slot.



4. Repeat Step 3 on the other bridges until all of the bridges have been cabled.

Secure or unsecure the FibreBridge bridge

To easily disable potentially unsecure Ethernet protocols on a bridge, beginning with ONTAP 9.5 you can secure the bridge. This disables the bridge's Ethernet ports. You can also reenable Ethernet access.

About this task

- Securing the bridge disables telnet and other IP port protocols and services (FTP, ExpressNAV, ICMP, or QuickNAV) on the bridge.
- This procedure uses out-of-band management using the ONTAP prompt, which is available beginning with ONTAP 9.5.

  You can issue the commands from the bridge CLI if you are not using out-of-band management.
- The `unsecurebridge` command can be used to re-enable the Ethernet ports.
- In ONTAP 9.7 and earlier, running the `securebridge` command on the ATTO FibreBridge might not update the bridge status correctly on the partner cluster. If this occurs, run the `securebridge` command from the partner cluster.

> Beginning with ONTAP 9.8, the `storage bridge` command is replaced with `system bridge`. The following steps show the `storage bridge` command, but if you are running ONTAP 9.8 or later, the `system bridge` command is preferred.

Steps

1. From the ONTAP prompt of the cluster containing the bridge, secure or unsecure the bridge.
   - The following command secures bridge_A_1:

     ```
     cluster_A> storage bridge run-cli -bridge bridge_A_1 -command securebridge
     ```

   - The following command unsecures bridge_A_1:

     ```
     cluster_A> storage bridge run-cli -bridge bridge_A_1 -command unsecurebridge
     ```

2. From the ONTAP prompt of the cluster containing the bridge, save the bridge configuration:

```
storage bridge run-cli -bridge <bridge-name> -command saveconfiguration
```

The following command secures bridge_A_1:

```
cluster_A> storage bridge run-cli -bridge bridge_A_1 -command
saveconfiguration
```

3. From the ONTAP prompt of the cluster containing the bridge, restart the bridge's firmware:

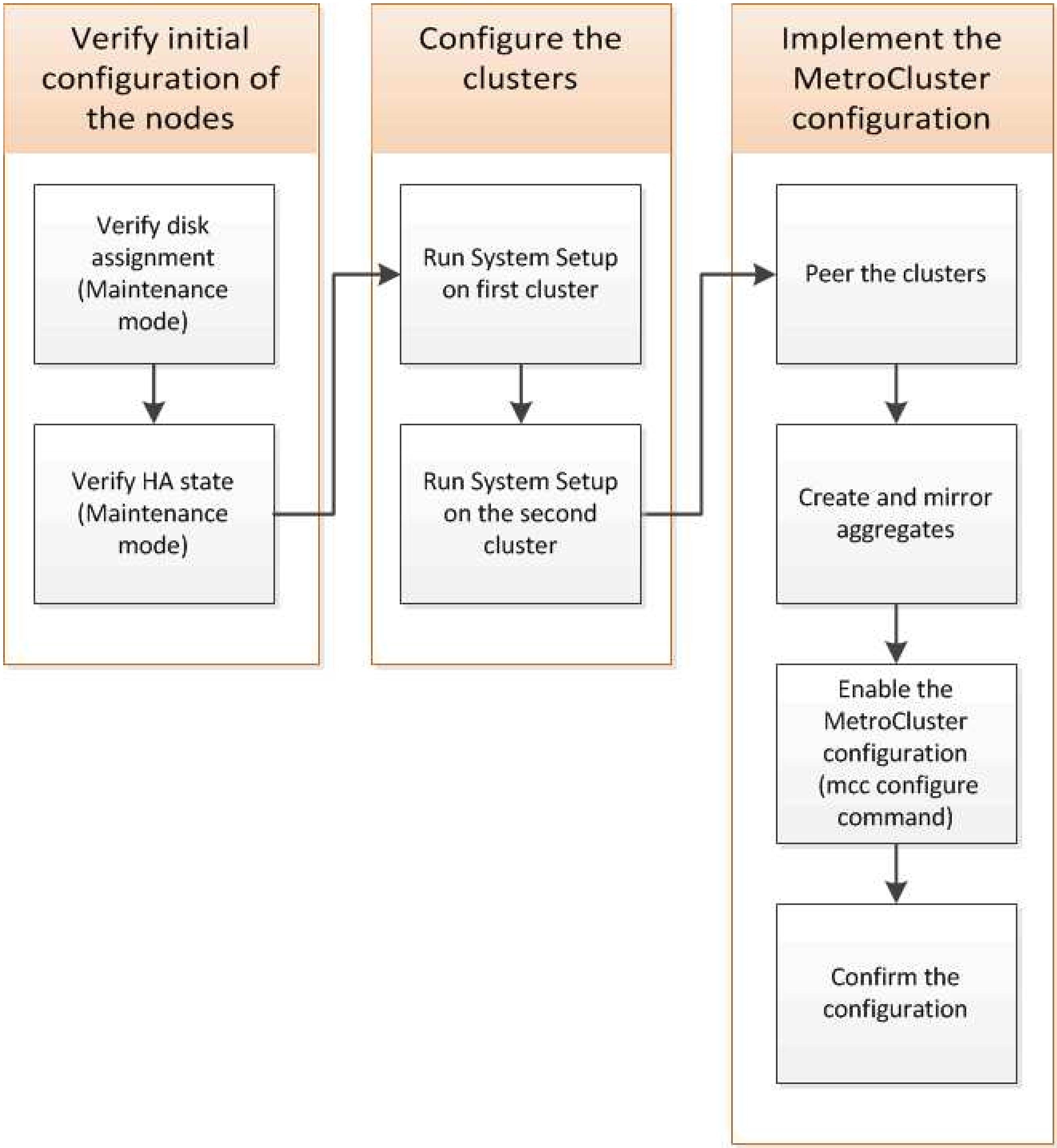
```
storage bridge run-cli -bridge <bridge-name> -command firmwarerestart
```

The following command secures bridge_A_1:

```
cluster_A> storage bridge run-cli -bridge bridge_A_1 -command firmwarerestart
```

Configuring the MetroCluster software in ONTAP

You must set up each node in the MetroCluster configuration in ONTAP, including the node-level configurations and the configuration of the nodes into two sites. You must also implement the MetroCluster relationship between the two sites.



Steps

1. Gather the required IP addresses for the controller modules before you begin the configuration process.
2. Complete the IP network information worksheet for site A.

IP network information worksheet for Site A

You must obtain IP addresses and other network information for the first MetroCluster site (site A) from your network administrator before you configure the system.

Site A cluster creation information

When you first create the cluster, you need the following information:

Type of information	Your values
Cluster name. Example used in this information: site_A	
DNS domain	
DNS name servers	
Location	
Administrator password	

Site A node information

For each node in the cluster, you need a management IP address, a network mask, and a default gateway.

Node	Port	IP address	Network mask	Default gateway
Node 1. Example used in this information: controller_A_1				
Node 2. Not required if using two-node MetroCluster configuration (one node at each site). Example used in this information: controller_A_2				

Site A LIFs and ports for cluster peering

For each node in the cluster, you need the IP addresses of two intercluster LIFs, including a network mask and a default gateway. The intercluster LIFs are used to peer the clusters.

Node	Port	IP address of intercluster LIF	Network mask	Default gateway
Node 1 IC LIF 1				
Node 1 IC LIF 2				

Site A time server information

You must synchronize the time, which requires one or more NTP time servers.

Node	Host name	IP address	Network mask	Default gateway
NTP server 1				
NTP server 2				

Site A AutoSupport information

You must configure AutoSupport on each node, which requires the following information:

Type of information		Your values
From email address		
Mail hosts	IP addresses or names	
Transport protocol	HTTP, HTTPS, or SMTP	
	Proxy server	
Recipient email addresses or distribution lists	Full-length messages	
	Concise messages	
	Partners	

Site A SP information

You must enable access to the Service Processor (SP) of each node for troubleshooting and maintenance. This requires the following network information for each node:

Node	IP address	Network mask	Default gateway
Node 1			

IP network information worksheet for site B

You must obtain IP addresses and other network information for the second MetroCluster site (site B) from your network administrator before you configure the system.

Site B cluster creation information

When you first create the cluster, you need the following information:

Type of information	Your values
Cluster name. Example used in this information: site_B	
DNS domain	
DNS name servers	
Location	
Administrator password	

Site B node information

For each node in the cluster, you need a management IP address, a network mask, and a default gateway.

Node	Port	IP address	Network mask	Default gateway
Node 1. Example used in this information: controller_B_1				
Node 2. Not required for two-node MetroCluster configurations (one node at each site). Example used in this information: controller_B_2				

Site B LIFs and ports for cluster peering

For each node in the cluster, you need the IP addresses of two intercluster LIFs, including a network mask and a default gateway. The intercluster LIFs are used to peer the clusters.

Node	Port	IP address of intercluster LIF	Network mask	Default gateway
Node 1 IC LIF 1				
Node 1 IC LIF 2				

Site B time server information

You must synchronize the time, which requires one or more NTP time servers.

Node	Host name	IP address	Network mask	Default gateway
NTP server 1				
NTP server 2				

Site B AutoSupport information

You must configure AutoSupport on each node, which requires the following information:

Type of information		Your values
From email address		
Mail hosts	IP addresses or names	
Transport protocol	HTTP, HTTPS, or SMTP	
	Proxy server	
Recipient email addresses or distribution lists	Full-length messages	
	Concise messages	
	Partners	

Site B SP information

You must enable access to the Service Processor (SP) of each node for troubleshooting and maintenance, which requires the following network information for each node:

Node	IP address	Network mask	Default gateway
Node 1 (controller_B_1)			

Similarities and differences between standard cluster and MetroCluster configurations

The configuration of the nodes in each cluster in a MetroCluster configuration is similar to that of nodes in a standard cluster.

The MetroCluster configuration is built on two standard clusters. Physically, the configuration must be symmetrical, with each node having the same hardware configuration, and all of the MetroCluster components must be cabled and configured. However, the basic software configuration for nodes in a MetroCluster configuration is the same as that for nodes in a standard cluster.

Configuration step	Standard cluster configuration	MetroCluster configuration

Configure management, cluster, and data LIFs on each node.	Same in both types of clusters	
Configure the root aggregate.	Same in both types of clusters	
Set up the cluster on one node in the cluster.	Same in both types of clusters	
Join the other node to the cluster.	Same in both types of clusters	
Create a mirrored root aggregate.	Optional	Required
Peer the clusters.	Optional	Required
Enable the MetroCluster configuration.	Does not apply	Required

Restoring system defaults and configuring the HBA type on a controller module

To ensure a successful MetroCluster installation, reset and restore defaults on the controller modules.

Important

This task is only required for stretch configurations using FC-to-SAS bridges.

Steps

1. At the LOADER prompt, return the environmental variables to their default setting:

   ```
   set-defaults
   ```

2. Boot the node into Maintenance mode, then configure the settings for any HBAs in the system:

   a. Boot into Maintenance mode:

      ```
      boot_ontap maint
      ```

   b. Check the current settings of the ports:

      ```
      ucadmin show
      ```

   c. Update the port settings as needed.

If you have this type of HBA and desired mode…	Use this command…
CNA FC	`ucadmin modify -m fc -t initiator adapter_name`
CNA Ethernet	`ucadmin modify -mode cna adapter_name`

FC target	`fcadmin config -t target` *`adapter_name`*
FC initiator	`fcadmin config -t initiator` *`adapter_name`*

3. Exit Maintenance mode:

   `halt`

   After you run the command, wait until the node stops at the LOADER prompt.

4. Boot the node back into Maintenance mode to enable the configuration changes to take effect:

   `boot_ontap maint`

5. Verify the changes you made:

If you have this type of HBA…	Use this command…
CNA	`ucadmin show`
FC	`fcadmin show`

6. Exit Maintenance mode:

   `halt`

   After you run the command, wait until the node stops at the LOADER prompt.

7. Boot the node to the boot menu:

   `boot_ontap menu`

   After you run the command, wait until the boot menu is shown.

8. Clear the node configuration by typing “wipeconfig” at the boot menu prompt, and then press Enter.

   The following screen shows the boot menu prompt:

```
Please choose one of the following:

     (1) Normal Boot.
     (2) Boot without /etc/rc.
     (3) Change password.
     (4) Clean configuration and initialize all disks.
     (5) Maintenance mode boot.
     (6) Update flash from backup config.
     (7) Install new software first.
     (8) Reboot node.
     (9) Configure Advanced Drive Partitioning.
     Selection (1-9)?  wipeconfig
 This option deletes critical system configuration, including cluster
membership.
 Warning: do not run this option on a HA node that has been taken over.
 Are you sure you want to continue?: yes
 Rebooting to finish wipeconfig request.
```

Configuring FC-VI ports on a X1132A-R6 quad-port card on FAS8020 systems

If you are using the X1132A-R6 quad-port card on a FAS8020 system, you can enter Maintenance mode to configure the 1a and 1b ports for FC-VI and initiator usage. This is not required on MetroCluster systems received from the factory, in which the ports are set appropriately for your configuration.

About this task

This task must be performed in Maintenance mode.

> Converting an FC port to an FC-VI port with the ucadmin command is only supported on the FAS8020 and AFF 8020 systems. Converting FC ports to FCVI ports is not supported on any other platform.

Steps

1. Disable the ports:

   ```
   storage disable adapter 1a
   ```

   ```
   storage disable adapter 1b
   ```

```
*> storage disable adapter 1a
Jun 03 02:17:57 [controller_B_1:fci.adapter.offlining:info]: Offlining
Fibre Channel adapter 1a.
Host adapter 1a disable succeeded
Jun 03 02:17:57 [controller_B_1:fci.adapter.offline:info]: Fibre Channel
adapter 1a is now offline.
*> storage disable adapter 1b
Jun 03 02:18:43 [controller_B_1:fci.adapter.offlining:info]: Offlining
Fibre Channel adapter 1b.
Host adapter 1b disable succeeded
Jun 03 02:18:43 [controller_B_1:fci.adapter.offline:info]: Fibre Channel
adapter 1b is now offline.
*>
```

2. Verify that the ports are disabled:

   `ucadmin show`

```
*> ucadmin show
         Current  Current    Pending  Pending    Admin
Adapter  Mode     Type       Mode     Type       Status
-------  -------  ---------  -------  ---------  -------
  ...
  1a     fc       initiator  -        -          offline
  1b     fc       initiator  -        -          offline
  1c     fc       initiator  -        -          online
  1d     fc       initiator  -        -          online
```

3. Set the a and b ports to FC-VI mode:

   `ucadmin modify -adapter 1a -type fcvi`

   The command sets the mode on both ports in the port pair, 1a and 1b (even though only 1a is specified in the command).

```
*> ucadmin modify -t fcvi 1a
Jun 03 02:19:13 [controller_B_1:ucm.type.changed:info]: FC-4 type has
changed to fcvi on adapter 1a. Reboot the controller for the changes to
take effect.
Jun 03 02:19:13 [controller_B_1:ucm.type.changed:info]: FC-4 type has
changed to fcvi on adapter 1b. Reboot the controller for the changes to
take effect.
```

4. Confirm that the change is pending:

ucadmin show

```
*> ucadmin show
         Current  Current    Pending  Pending    Admin
Adapter  Mode     Type       Mode     Type       Status
-------  -------  ---------  -------  ---------  -------
  ...
  1a     fc       initiator  -        fcvi       offline
  1b     fc       initiator  -        fcvi       offline
  1c     fc       initiator  -        -          online
  1d     fc       initiator  -        -          online
```

5. Shut down the controller, and then reboot into Maintenance mode.
6. Confirm the configuration change:

ucadmin show local

```
Node            Adapter  Mode     Type       Mode     Type       Status
------------    -------  -------  ---------  -------  ---------
-----------
...
controller_B_1
                1a       fc       fcvi       -        -          online
controller_B_1
                1b       fc       fcvi       -        -          online
controller_B_1
                1c       fc       initiator  -        -          online
controller_B_1
                1d       fc       initiator  -        -          online
6 entries were displayed.
```

Verifying disk assignment in Maintenance mode in a two-node configuration

Before fully booting the system to ONTAP, you can optionally boot the system to Maintenance mode and verify the disk assignment on the nodes. The disks should be assigned to create a fully symmetric configuration with both sites owning their own disk shelves and serving data, where each node and each pool have an equal number of mirrored disks assigned to them.

Before you begin

The system must be in Maintenance mode.

About this task

New MetroCluster systems have disk assignments completed prior to shipment.

The following table shows example pool assignments for a MetroCluster configuration. Disks are assigned to pools on a per-shelf basis.

Disk shelf (*example name*)…	**At site…**	**Belongs to…**	**And is assigned to that node's…**
Disk shelf 1 (shelf_A_1_1)	Site A	Node A 1	Pool 0
Disk shelf 2 (shelf_A_1_3)			
Disk shelf 3 (shelf_B_1_1)		Node B 1	Pool 1
Disk shelf 4 (shelf_B_1_3)			
Disk shelf 9 (shelf_B_1_2)	Site B	Node B 1	Pool 0
Disk shelf 10 (shelf_B_1_4)			
Disk shelf 11 (shelf_A_1_2)		Node A 1	Pool 1
Disk shelf 12 (shelf_A_1_4)			

If your configuration includes DS460C disk shelves, you should manually assign the disks using the following guidelines for each 12-disk drawer:

Assign these disks in the drawer…	**To this node and pool…**
1 - 6	Local node's pool 0
7 - 12	DR partner's pool 1

This disk assignment pattern minimizes the effect on an aggregate if a drawer goes offline.

Steps

1. If your system was received from the factory, confirm the shelf assignments:

   ```
   disk show –v
   ```

2. If necessary, you can explicitly assign disks on the attached disk shelves to the appropriate pool

   ```
   disk assign
   ```

   Disk shelves at the same site as the node are assigned to pool 0 and disk shelves located at the partner site are assigned to pool 1. You should assign an equal number of shelves to each pool.

   a. If you have not done so, boot each system into Maintenance mode.

   b. On the node on site A, systematically assign the local disk shelves to pool 0 and the remote disk shelves to pool 1:
   `disk assign -shelf *disk_shelf_name* -p *pool*`

If storage controller node_A_1 has four shelves, you issue the following commands:

```
*> disk assign -shelf shelf_A_1_1 -p 0
*> disk assign -shelf shelf_A_1_3 -p 0

*> disk assign -shelf shelf_A_1_2 -p 1
*> disk assign -shelf shelf_A_1_4 -p 1
```

c. On the node at the remote site (site B), systematically assign its local disk shelves to pool 0 and its remote disk shelves to pool 1:
`disk assign -shelf` *`disk_shelf_name`* `-p` *`pool`*

If storage controller node_B_1 has four shelves, you issue the following commands:

```
*> disk assign -shelf shelf_B_1_2   -p 0
*> disk assign -shelf shelf_B_1_4  -p 0

*> disk assign -shelf shelf_B_1_1 -p 1
 *> disk assign -shelf shelf_B_1_3 -p 1
```

d. Show the disk shelf IDs and bays for each disk:
`disk show –v`

Verifying the HA state of components

In a stretch MetroCluster configuration that is not preconfigured at the factory, you must verify that the HA state of the controller and chassis component is set to "mcc-2n" so that they boot up properly. For systems received from the factory, this value is preconfigured and you do not need to verify it.

Before you begin

The system must be in Maintenance mode.

Steps

1. In Maintenance mode, view the HA state of the controller module and chassis:

   `ha-config show`

   The controller module and chassis should show the value "mcc-2n".

2. If the displayed system state of the controller is not "mcc-2n", set the HA state for the controller:

   `ha-config modify controller mcc-2n`

3. If the displayed system state of the chassis is not "mcc-2n", set the HA state for the chassis:

   `ha-config modify chassis mcc-2n`

   Halt the node.

Wait until the node is back at the LOADER prompt.

4. Repeat these steps on each node in the MetroCluster configuration.

Setting up ONTAP in a two-node MetroCluster configuration

In a two-node MetroCluster configuration, on each cluster you must boot up the node, exit the Cluster Setup wizard, and use the `cluster setup` command to configure the node into a single-node cluster.

Before you begin

You must not have configured the Service Processor.

About this task

This task is for two-node MetroCluster configurations using native NetApp storage.

This task must be performed on both clusters in the MetroCluster configuration.

For more general information about setting up ONTAP, see the Setup ONTAP

Steps

1. Power on the first node.

   > ⓘ You must repeat this step on the node at the disaster recovery (DR) site.

   The node boots, then the Cluster Setup wizard starts on the console informing you that AutoSupport will be enabled automatically.

```
::> Welcome to the cluster setup wizard.

You can enter the following commands at any time:
  "help" or "?" - if you want to have a question clarified,
  "back" - if you want to change previously answered questions, and
  "exit" or "quit" - if you want to quit the cluster setup wizard.
     Any changes you made before quitting will be saved.

You can return to cluster setup at any time by typing "cluster setup".
To accept a default or omit a question, do not enter a value.

This system will send event messages and periodic reports to NetApp
Technical
Support. To disable this feature, enter
autosupport modify -support disable
within 24 hours.

Enabling AutoSupport can significantly speed problem determination and
resolution, should a problem occur on your system.
For further information on AutoSupport, see:
http://support.netapp.com/autosupport/

Type yes to confirm and continue {yes}: yes

Enter the node management interface port [e0M]:
Enter the node management interface IP address [10.101.01.01]:

Enter the node management interface netmask [101.010.101.0]:
Enter the node management interface default gateway [10.101.01.0]:



Do you want to create a new cluster or join an existing cluster?
{create, join}:
```

2. Create a new cluster:

   `create`

3. Choose whether the node is to be used as a single node cluster.

```
Do you intend for this node to be used as a single node cluster? {yes,
no} [yes]:
```

4. Accept the system default “yes” by pressing Enter, or enter your own values by typing “no”, and then

pressing Enter.

5. Follow the prompts to complete the **Cluster Setup** wizard, pressing Enter to accept the default values or typing your own values and then pressing Enter.

   The default values are determined automatically based on your platform and network configuration.

6. After you complete the **Cluster Setup** wizard and it exits, verify that the cluster is active and the first node is healthy:

   `cluster show`

   The following example shows a cluster in which the first node (cluster1-01) is healthy and eligible to participate:

```
cluster1::> cluster show
Node                  Health  Eligibility
--------------------- ------- ------------
cluster1-01           true    true
```

   If it becomes necessary to change any of the settings you entered for the admin SVM or node SVM, you can access the **Cluster Setup** wizard by using the `cluster setup` command.

Configuring the clusters into a MetroCluster configuration

You must peer the clusters, mirror the root aggregates, create a mirrored data aggregate, and then issue the command to implement the MetroCluster operations.

Peering the clusters

The clusters in the MetroCluster configuration must be in a peer relationship so that they can communicate with each other and perform the data mirroring essential to MetroCluster disaster recovery.

Related information

Cluster and SVM peering express configuration

Considerations when using dedicated ports

Considerations when sharing data ports

Configuring intercluster LIFs

You must create intercluster LIFs on ports used for communication between the MetroCluster partner clusters. You can use dedicated ports or ports that also have data traffic.

Configuring intercluster LIFs on dedicated ports

You can configure intercluster LIFs on dedicated ports. Doing so typically increases the available bandwidth for replication traffic.

Steps

1. List the ports in the cluster:

```
network port show
```

For complete command syntax, see the man page.

The following example shows the network ports in “cluster01”:

```
cluster01::> network port show
                                                             Speed
(Mbps)
Node   Port      IPspace      Broadcast Domain Link   MTU    Admin/Oper
------ --------- ------------ ---------------- ----- -------
------------
cluster01-01
       e0a       Cluster      Cluster          up     1500   auto/1000
       e0b       Cluster      Cluster          up     1500   auto/1000
       e0c       Default      Default          up     1500   auto/1000
       e0d       Default      Default          up     1500   auto/1000
       e0e       Default      Default          up     1500   auto/1000
       e0f       Default      Default          up     1500   auto/1000
cluster01-02
       e0a       Cluster      Cluster          up     1500   auto/1000
       e0b       Cluster      Cluster          up     1500   auto/1000
       e0c       Default      Default          up     1500   auto/1000
       e0d       Default      Default          up     1500   auto/1000
       e0e       Default      Default          up     1500   auto/1000
       e0f       Default      Default          up     1500   auto/1000
```

2. Determine which ports are available to dedicate to intercluster communication:

   ```
   network interface show -fields home-port,curr-port
   ```

   For complete command syntax, see the man page.

   The following example shows that ports “e0e” and “e0f” have not been assigned LIFs:

```
cluster01::> network interface show -fields home-port,curr-port
vserver lif                  home-port curr-port

Cluster cluster01-01_clus1   e0a       e0a
Cluster cluster01-01_clus2   e0b       e0b
Cluster cluster01-02_clus1   e0a       e0a
Cluster cluster01-02_clus2   e0b       e0b
cluster01
        cluster_mgmt         e0c       e0c
cluster01
        cluster01-01_mgmt1   e0c       e0c
cluster01
        cluster01-02_mgmt1   e0c       e0c
```

3. Create a failover group for the dedicated ports:

   `network interface failover-groups create -vserver` *`system_SVM`* `-failover-group` *`failover_group`* `-targets` *`physical_or_logical_ports`*

   The following example assigns ports “e0e” and “e0f” to the failover group “intercluster01” on system SVM “cluster01”:

```
cluster01::> network interface failover-groups create -vserver cluster01
-failover-group
intercluster01 -targets
cluster01-01:e0e,cluster01-01:e0f,cluster01-02:e0e,cluster01-02:e0f
```

4. Verify that the failover group was created:

   `network interface failover-groups show`

   For complete command syntax, see the man page.

```
cluster01::> network interface failover-groups show
                                  Failover
Vserver           Group           Targets
---------------- ----------------
--------------------------------------------
Cluster
                  Cluster
                                  cluster01-01:e0a, cluster01-01:e0b,
                                  cluster01-02:e0a, cluster01-02:e0b
cluster01
                  Default
                                  cluster01-01:e0c, cluster01-01:e0d,
                                  cluster01-02:e0c, cluster01-02:e0d,
                                  cluster01-01:e0e, cluster01-01:e0f
                                  cluster01-02:e0e, cluster01-02:e0f
                  intercluster01
                                  cluster01-01:e0e, cluster01-01:e0f
                                  cluster01-02:e0e, cluster01-02:e0f
```

5. Create intercluster LIFs on the system SVM and assign them to the failover group.

ONTAP version	Command
ONTAP 9.6 and later	`network interface create -vserver system_SVM -lif LIF_name -service-policy default-intercluster -home -node node -home-port port -address port_IP -netmask netmask -failover-group failover_group`
ONTAP 9.5 and earlier	`network interface create -vserver system_SVM -lif LIF_name -role intercluster -home-node node -home-port port -address port_IP -netmask netmask -failover-group failover_group`

For complete command syntax, see the man page.

The following example creates intercluster LIFs “cluster01_icl01” and “cluster01_icl02” in the failover group “intercluster01”:

```
cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl01 -service-
policy default-intercluster -home-node cluster01-01 -home-port e0e
-address 192.168.1.201
-netmask 255.255.255.0 -failover-group intercluster01

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl02 -service-
policy default-intercluster -home-node cluster01-02 -home-port e0e
-address 192.168.1.202
-netmask 255.255.255.0 -failover-group intercluster01
```

6. Verify that the intercluster LIFs were created:

ONTAP version	Command
ONTAP 9.6 and later	`network interface show -service-policy default-intercluster`
ONTAP 9.5 and earlier	`network interface show -role intercluster`

For complete command syntax, see the man page.

```
cluster01::> network interface show -service-policy default-intercluster
            Logical    Status     Network            Current
Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node          Port
Home
----------- ---------- ---------- ------------------ -------------
------- ----
cluster01
            cluster01_icl01
                       up/up      192.168.1.201/24   cluster01-01  e0e
true
            cluster01_icl02
                       up/up      192.168.1.202/24   cluster01-02  e0f
true
```

7. Verify that the intercluster LIFs are redundant:

ONTAP version	Command
ONTAP 9.6 and later	`network interface show -service-policy default-intercluster -failover`

In ONTAP 9.5 and earlier	`network interface show -role intercluster -failover`

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs “cluster01_icl01” and “cluster01_icl02” on the SVM port “e0e” will fail over to port “e0f”.

```
cluster01::> network interface show -service-policy default-intercluster
-failover
         Logical         Home                  Failover        Failover
Vserver  Interface       Node:Port             Policy          Group
-------- --------------- --------------------- --------------- --------
cluster01
         cluster01_icl01 cluster01-01:e0e   local-only
intercluster01
                            Failover Targets:  cluster01-01:e0e,
                                               cluster01-01:e0f
         cluster01_icl02 cluster01-02:e0e   local-only
intercluster01
                            Failover Targets:  cluster01-02:e0e,
                                               cluster01-02:e0f
```

Related information

Considerations when using dedicated ports

Configuring intercluster LIFs on shared data ports

You can configure intercluster LIFs on ports shared with the data network. Doing so reduces the number of ports you need for intercluster networking.

Steps

1. List the ports in the cluster:

   `network port show`

   For complete command syntax, see the man page.

   The following example shows the network ports in “cluster01”:

```
cluster01::> network port show
                                                             Speed
(Mbps)
Node   Port      IPspace      Broadcast Domain Link   MTU    Admin/Oper
------ --------- ------------ ---------------- ----- -------
------------
cluster01-01
       e0a       Cluster      Cluster          up     1500   auto/1000
       e0b       Cluster      Cluster          up     1500   auto/1000
       e0c       Default      Default          up     1500   auto/1000
       e0d       Default      Default          up     1500   auto/1000
cluster01-02
       e0a       Cluster      Cluster          up     1500   auto/1000
       e0b       Cluster      Cluster          up     1500   auto/1000
       e0c       Default      Default          up     1500   auto/1000
       e0d       Default      Default          up     1500   auto/1000
```

2. Create intercluster LIFs on the system SVM:

ONTAP version	Command
ONTAP 9.6 and later	`network interface create -vserver` *`system_SVM`* `-lif` *`LIF_name`* `-service-policy default-intercluster -home -node node -home-port` *`port`* `-address` *`port_IP`* `-netmask` *`netmask`*
ONTAP 9.5 and earlier	`network interface create -vserver` *`system_SVM`* `-lif` *`LIF_name`* `-role intercluster -home-node` *`node`* `-home-port` *`port`* `-address` *`port_IP`* `-netmask` *`netmask`*

For complete command syntax, see the man page.

The following example creates intercluster LIFs “cluster01_icl01” and “cluster01_icl02”:

```
cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl01 -service-
policy default-intercluster -home-node cluster01-01 -home-port e0c
-address 192.168.1.201
-netmask 255.255.255.0

cluster01::> network interface create -vserver cluster01 -lif
cluster01_icl02 -service-
policy default-intercluster -home-node cluster01-02 -home-port e0c
-address 192.168.1.202
-netmask 255.255.255.0
```

3. Verify that the intercluster LIFs were created:

ONTAP version	Command
ONTAP 9.6 and later	`network interface show -service-policy default-intercluster`
ONTAP 9.5 and earlier	`network interface show -role intercluster`

For complete command syntax, see the man page.

```
cluster01::> network interface show -service-policy default-intercluster
            Logical    Status     Network            Current
Current Is
Vserver     Interface  Admin/Oper Address/Mask       Node          Port
Home
----------- ---------- ---------- ------------------ -------------
------- ----
cluster01
            cluster01_icl01
                       up/up      192.168.1.201/24   cluster01-01  e0c
true
            cluster01_icl02
                       up/up      192.168.1.202/24   cluster01-02  e0c
true
```

4. Verify that the intercluster LIFs are redundant:

ONTAP version	Command
ONTAP 9.6 and later	`network interface show –service-policy default-intercluster -failover`
ONTAP 9.5 and earlier	`network interface show -role intercluster -failover`

For complete command syntax, see the man page.

The following example shows that the intercluster LIFs “cluster01_icl01” and “cluster01_icl02” on port “e0c” will fail over to port “e0d”.

```
cluster01::> network interface show -service-policy default-intercluster
-failover
         Logical         Home                  Failover        Failover
Vserver  Interface       Node:Port             Policy          Group
-------- --------------- --------------------- --------------- --------
cluster01
         cluster01_icl01 cluster01-01:e0c   local-only
192.168.1.201/24
                            Failover Targets: cluster01-01:e0c,
                                              cluster01-01:e0d
         cluster01_icl02 cluster01-02:e0c   local-only
192.168.1.201/24
                            Failover Targets: cluster01-02:e0c,
                                              cluster01-02:e0d
```

Related information

Considerations when sharing data ports

Creating a cluster peer relationship

You must create the cluster peer relationship between the MetroCluster clusters.

Creating a cluster peer relationship

You can use the `cluster peer create` command to create a peer relationship between a local and remote cluster. After the peer relationship has been created, you can run `cluster peer create` on the remote cluster to authenticate it to the local cluster.

Before you begin

- You must have created intercluster LIFs on every node in the clusters that are being peered.
- The clusters must be running ONTAP 9.3 or later.

Steps

1. On the destination cluster, create a peer relationship with the source cluster:

   `cluster peer create -generate-passphrase -offer-expiration` *`MM/DD/YYYY HH:MM:SS`*`|1…7days|1…168hours -peer-addrs` *`peer_LIF_IPs`* `-ipspace` *`ipspace`*

   If you specify both `-generate-passphrase` and `-peer-addrs`, only the cluster whose intercluster LIFs are specified in `-peer-addrs` can use the generated password.

   You can ignore the `-ipspace` option if you are not using a custom IPspace. For complete command syntax, see the man page.

   The following example creates a cluster peer relationship on an unspecified remote cluster:

```
cluster02::> cluster peer create -generate-passphrase -offer-expiration
2days

                     Passphrase: UCa+6lRVICXeL/gq1WrK7ShR
                Expiration Time: 6/7/2017 08:16:10 EST
  Initial Allowed Vserver Peers: -
            Intercluster LIF IP: 192.140.112.101
              Peer Cluster Name: Clus_7ShR (temporary generated)

Warning: make a note of the passphrase - it cannot be displayed again.
```

2. On source cluster, authenticate the source cluster to the destination cluster:

   `cluster peer create -peer-addrs peer_LIF_IPs -ipspace ipspace`

   For complete command syntax, see the man page.

   The following example authenticates the local cluster to the remote cluster at intercluster LIF IP addresses 192.140.112.101 and 192.140.112.102:

```
cluster01::> cluster peer create -peer-addrs
192.140.112.101,192.140.112.102

Notice: Use a generated passphrase or choose a passphrase of 8 or more
characters.
        To ensure the authenticity of the peering relationship, use a
phrase or sequence of characters that would be hard to guess.

Enter the passphrase:
Confirm the passphrase:

Clusters cluster02 and cluster01 are peered.
```

   Enter the passphrase for the peer relationship when prompted.

3. Verify that the cluster peer relationship was created:

   `cluster peer show -instance`

```
cluster01::> cluster peer show -instance

                               Peer Cluster Name: cluster02
                   Remote Intercluster Addresses: 192.140.112.101,
192.140.112.102
              Availability of the Remote Cluster: Available
                             Remote Cluster Name: cluster2
                             Active IP Addresses: 192.140.112.101,
192.140.112.102
                           Cluster Serial Number: 1-80-123456
                  Address Family of Relationship: ipv4
            Authentication Status Administrative: no-authentication
               Authentication Status Operational: absent
                                Last Update Time: 02/05 21:05:41
                    IPspace for the Relationship: Default
```

4. Check the connectivity and status of the nodes in the peer relationship:

`cluster peer health show`

```
cluster01::> cluster peer health show
Node       cluster-Name                Node-Name
             Ping-Status               RDB-Health Cluster-Health  Avail…
---------- --------------------------- ---------  ---------------
--------
cluster01-01
           cluster02                   cluster02-01
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
                                       cluster02-02
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
cluster01-02
           cluster02                   cluster02-01
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
                                       cluster02-02
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
```

Creating a cluster peer relationship (ONTAP 9.2 and earlier)

You can use the `cluster peer create` command to initiate a request for a peering relationship between a local and remote cluster. After the peer relationship has been requested by the local cluster, you can run `cluster peer create` on the remote cluster to accept the relationship.

Before you begin

- You must have created intercluster LIFs on every node in the clusters being peered.
- The cluster administrators must have agreed on the passphrase each cluster will use to authenticate itself to the other.

Steps

1. On the data protection destination cluster, create a peer relationship with the data protection source cluster:

   `cluster peer create -peer-addrs` *`peer_LIF_IPs`* `-ipspace` *`ipspace`*

   You can ignore the `-ipspace` option if you are not using a custom IPspace. For complete command syntax, see the man page.

   The following example creates a cluster peer relationship with the remote cluster at intercluster LIF IP addresses 192.168.2.201 and 192.168.2.202:

   ```
   cluster02::> cluster peer create -peer-addrs 192.168.2.201,192.168.2.202
   Enter the passphrase:
   Please enter the passphrase again:
   ```

   Enter the passphrase for the peer relationship when prompted.

2. On the data protection source cluster, authenticate the source cluster to the destination cluster:

   `cluster peer create -peer-addrs` *`peer_LIF_IPs`* `-ipspace` *`ipspace`*

   For complete command syntax, see the man page.

   The following example authenticates the local cluster to the remote cluster at intercluster LIF IP addresses 192.140.112.203 and 192.140.112.204:

   ```
   cluster01::> cluster peer create -peer-addrs 192.168.2.203,192.168.2.204
   Please confirm the passphrase:
   Please confirm the passphrase again:
   ```

   Enter the passphrase for the peer relationship when prompted.

3. Verify that the cluster peer relationship was created:

   `cluster peer show –instance`

   For complete command syntax, see the man page.

```
cluster01::> cluster peer show -instance
Peer Cluster Name: cluster01
Remote Intercluster Addresses: 192.168.2.201,192.168.2.202
Availability: Available
Remote Cluster Name: cluster02
Active IP Addresses: 192.168.2.201,192.168.2.202
Cluster Serial Number: 1-80-000013
```

4. Check the connectivity and status of the nodes in the peer relationship:

   `cluster peer health show`

   For complete command syntax, see the man page.

```
cluster01::> cluster peer health show
Node       cluster-Name                Node-Name
             Ping-Status               RDB-Health Cluster-Health  Avail…
---------- --------------------------- ---------  ---------------
--------
cluster01-01
           cluster02                   cluster02-01
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
                                       cluster02-02
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
cluster01-02
           cluster02                   cluster02-01
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
                                       cluster02-02
             Data: interface_reachable
             ICMP: interface_reachable true       true            true
```

Mirroring the root aggregates

You must mirror the root aggregates to provide data protection.

About this task

By default, the root aggregate is created as RAID-DP type aggregate. You can change the root aggregate from RAID-DP to RAID4 type aggregate. The following command modifies the root aggregate for RAID4 type aggregate:

`storage aggregate modify -aggregate aggr_name -raidtype raid4`

On non-ADP systems, the RAID type of the aggregate can be modified from the default RAID-DP to RAID4 before or after the aggregate is mirrored.

Steps

1. Mirror the root aggregate:

   `storage aggregate mirror aggr_name`

   The following command mirrors the root aggregate for “controller_A_1”:

   ```
   controller_A_1::> storage aggregate mirror aggr0_controller_A_1
   ```

   This mirrors the aggregate, so it consists of a local plex and a remote plex located at the remote MetroCluster site.

2. Repeat the previous step for each node in the MetroCluster configuration.

Related information

Logical storage management

ONTAP concepts

Creating a mirrored data aggregate on each node

You must create a mirrored data aggregate on each node in the DR group.

Before you begin

- You should know what drives will be used in the new aggregate.
- If you have multiple drive types in your system (heterogeneous storage), you should understand how you can ensure that the correct drive type is selected.

About this task

- Drives are owned by a specific node; when you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.

  Disk and aggregate management

- Aggregate names must be unique across the MetroCluster sites. This means that you cannot have two different aggregates with the same name on site A and site B.

Steps

1. Display a list of available spares:

   `storage disk show -spare -owner node_name`

2. Create the aggregate:

   `storage aggregate create -mirror true`

If you are logged in to the cluster on the cluster management interface, you can create an aggregate on any node in the cluster. To ensure that the aggregate is created on a specific node, use the `-node` parameter or specify drives that are owned by that node.

You can specify the following options:

- Aggregate's home node (that is, the node that owns the aggregate in normal operation)
- List of specific drives that are to be added to the aggregate
- Number of drives to include

  > In the minimum supported configuration, in which a limited number of drives are available, you must use the force-small-aggregate option to allow the creation of a three disk RAID-DP aggregate.

- Checksum style to use for the aggregate
- Type of drives to use
- Size of drives to use
- Drive speed to use
- RAID type for RAID groups on the aggregate
- Maximum number of drives that can be included in a RAID group
- Whether drives with different RPM are allowed For more information about these options, see the `storage aggregate create` man page.

  The following command creates a mirrored aggregate with 10 disks:

  ```
  cluster_A::> storage aggregate create aggr1_node_A_1 -diskcount 10
  -node node_A_1 -mirror true
  [Job 15] Job is queued: Create aggr1_node_A_1.
  [Job 15] The job is starting.
  [Job 15] Job succeeded: DONE
  ```

3. Verify the RAID group and drives of your new aggregate:

   `storage aggregate show-status -aggregate` *`aggregate-name`*

Creating unmirrored data aggregates

You can optionally create unmirrored data aggregates for data that does not require the redundant mirroring provided by MetroCluster configurations.

Before you begin

- You should know what drives will be used in the new aggregate.
- If you have multiple drive types in your system (heterogeneous storage), you should understand how you can verify that the correct drive type is selected.

Example 1. About this task

ATTENTION: In MetroCluster FC configurations, the unmirrored aggregates will only be online after a switchover if the remote disks in the aggregate are accessible. If the ISLs fail, the local node may be unable to access the data in the unmirrored remote disks. The failure of an aggregate can lead to a reboot of the local node.

> The unmirrored aggregates must be local to the node owning them.

- Drives are owned by a specific node; when you create an aggregate, all drives in that aggregate must be owned by the same node, which becomes the home node for that aggregate.
- Aggregate names should conform to the naming scheme you determined when you planned your MetroCluster configuration.
- The Disks and aggregates management contains more information about mirroring aggregates.

Steps

1. Display a list of available spares:

   `storage disk show -spare -owner` *`node_name`*

2. Create the aggregate:

   `storage aggregate create`

   If you are logged in to the cluster on the cluster management interface, you can create an aggregate on any node in the cluster. To verify that the aggregate is created on a specific node, you should use the `-node` parameter or specify drives that are owned by that node.

   You can specify the following options:

   - Aggregate's home node (that is, the node that owns the aggregate in normal operation)
   - List of specific drives that are to be added to the aggregate
   - Number of drives to include
   - Checksum style to use for the aggregate
   - Type of drives to use
   - Size of drives to use
   - Drive speed to use
   - RAID type for RAID groups on the aggregate
   - Maximum number of drives that can be included in a RAID group
   - Whether drives with different RPM are allowed For more information about these options, see the `storage aggregate create` man page.

     The following command creates a unmirrored aggregate with 10 disks:

```
controller_A_1::> storage aggregate create aggr1_controller_A_1
-diskcount 10 -node controller_A_1
[Job 15] Job is queued: Create aggr1_controller_A_1.
[Job 15] The job is starting.
[Job 15] Job succeeded: DONE
```

3. Verify the RAID group and drives of your new aggregate:

   `storage aggregate show-status -aggregate aggregate-name`

Implementing the MetroCluster configuration

You must run the `metrocluster configure` command to start data protection in a MetroCluster configuration.

Before you begin

- There should be at least two non-root mirrored data aggregates on each cluster.

  Additional data aggregates can be either mirrored or unmirrored.

  Verify the aggregate types:

  `storage aggregate show`

  > ⓘ If you want to use a single mirrored data aggregate, then see Configure MCC software in ONTAP for instructions.

- The ha-config state of the controllers and chassis must be “mcc-2n”.

About this task

You can issue the `metrocluster configure` command once, on any of the nodes, to enable the MetroCluster configuration. You do not need to issue the command on each of the sites or nodes, and it does not matter which node or site you choose to issue the command on.

Steps

1. Configure the MetroCluster in the following format:

If your MetroCluster configuration has…	Then do this…
Multiple data aggregates	From any node’s prompt, configure MetroCluster: `metrocluster configure node-name`

A single mirrored data aggregate	a. From any node's prompt, change to the advanced privilege level:
	
	`set -privilege advanced`
	
	You need to respond with "y" when you are prompted to continue into advanced mode and you see the advanced mode prompt (*>).
	
	b. Configure the MetroCluster with the `-allow-with-one -aggregate true` parameter:
	
	`metrocluster configure -allow-with-one-aggregate true` *`node-name`*
	
	c. Return to the admin privilege level:
	`set -privilege admin`

The best practice is to have multiple data aggregates. If the first DR group has only one aggregate and you want to add a DR group with one aggregate, you must move the metadata volume off the single data aggregate. For more information on this procedure, see Moving a metadata volume in MetroCluster configurations.

The following command enables the MetroCluster configuration on all of the nodes in the DR group that contains "controller_A_1":

```
cluster_A::*> metrocluster configure -node-name controller_A_1

[Job 121] Job succeeded: Configure is successful.
```

2. Verify the networking status on site A:

   `network port show`

   The following example shows the network port usage:

```
cluster_A::> network port show
                                                          Speed (Mbps)
Node   Port      IPspace   Broadcast Domain Link   MTU    Admin/Oper
------ --------- --------- ---------------- ----- ------- ------------
controller_A_1
       e0a       Cluster   Cluster          up     9000  auto/1000
       e0b       Cluster   Cluster          up     9000  auto/1000
       e0c       Default   Default          up     1500  auto/1000
       e0d       Default   Default          up     1500  auto/1000
       e0e       Default   Default          up     1500  auto/1000
       e0f       Default   Default          up     1500  auto/1000
       e0g       Default   Default          up     1500  auto/1000

7 entries were displayed.
```

3. Verify the MetroCluster configuration from both sites in the MetroCluster configuration.

   a. Verify the configuration from site A:
      `metrocluster show`

```
cluster_A::> metrocluster show

Cluster                    Entry Name          State
-------------------------- ------------------- -----------
 Local: cluster_A          Configuration state configured
                           Mode                normal
                           AUSO Failure Domain auso-on-cluster-
disaster
Remote: cluster_B          Configuration state configured
                           Mode                normal
                           AUSO Failure Domain auso-on-cluster-
disaster
```

   b. Verify the configuration from site B:
      `metrocluster show`

```
cluster_B::> metrocluster show
Cluster                    Entry Name          State
-------------------------- ------------------- -----------
 Local: cluster_B          Configuration state configured
                           Mode                normal
                           AUSO Failure Domain auso-on-cluster-
disaster
Remote: cluster_A          Configuration state configured
                           Mode                normal
                           AUSO Failure Domain auso-on-cluster-
disaster
```

Configuring FC-to-SAS bridges for health monitoring

In systems running ONTAP versions prior to 9.8, if your configuration includes FC-to-SAS bridges, you must perform some special configuration steps to monitor the FC-to-SAS bridges in the MetroCluster configuration.

- Third-party SNMP monitoring tools are not supported for FibreBridge bridges.
- Beginning with ONTAP 9.8, FC-to-SAS bridges are monitored via in-band connections by default, and additional configuration is not required.

> Beginning with ONTAP 9.8, the `storage bridge` command is replaced with `system bridge`. The following steps show the `storage bridge` command, but if you are running ONTAP 9.8 or later, the `system bridge` command is preferred.

Steps

1. From the ONTAP cluster prompt, add the bridge to health monitoring:
   a. Add the bridge, using the command for your version of ONTAP:

ONTAP version	Command
ONTAP 9.5 and later	`storage bridge add -address 0.0.0.0 -managed-by in-band -name` *`bridge-name`*
ONTAP 9.4 and earlier	`storage bridge add -address` *`bridge-ip-address`* `-name` *`bridge-name`*

   b. Verify that the bridge has been added and is properly configured:

   `storage bridge show`

   It might take as long as 15 minutes to reflect all of the data because of the polling interval. The ONTAP health monitor can contact and monitor the bridge if the value in the “Status” column is “ok”, and other information, such as the worldwide name (WWN), is displayed.

   The following example shows that the FC-to-SAS bridges are configured:

```
controller_A_1::> storage bridge show

Bridge              Symbolic Name Is Monitored  Monitor Status
Vendor Model                Bridge WWN
------------------  ------------- ------------  --------------
------ -----------------    ----------
ATTO_10.10.20.10  atto01        true          ok             Atto
FibreBridge 7500N      20000010867038c0
ATTO_10.10.20.11  atto02        true          ok             Atto
FibreBridge 7500N      20000010867033c0
ATTO_10.10.20.12  atto03        true          ok             Atto
FibreBridge 7500N      20000010867030c0
ATTO_10.10.20.13  atto04        true          ok             Atto
FibreBridge 7500N      2000001086703b80

4 entries were displayed

 controller_A_1::>
```

Checking the MetroCluster configuration

You can check that the components and relationships in the MetroCluster configuration are working correctly. You should do a check after initial configuration and after making any changes to the MetroCluster configuration. You should also do a check before a negotiated (planned) switchover or a switchback operation.

If the `metrocluster check run` command is issued twice within a short time on either or both clusters, a conflict can occur and the command might not collect all data. Subsequent `metrocluster check show` commands do not show the expected output.

1. Check the configuration:

   `metrocluster check run`

   The command runs as a background job and might not be completed immediately.

```
cluster_A::> metrocluster check run
The operation has been started and is running in the background. Wait
for
it to complete and run "metrocluster check show" to view the results. To
check the status of the running metrocluster check operation, use the
command,
"metrocluster operation history show -job-id 2245"
```

```
cluster_A::> metrocluster check show

Component           Result
------------------- ---------
nodes               ok
lifs                ok
config-replication  ok
aggregates          ok
clusters            ok
connections         ok
volumes             ok
7 entries were displayed.
```

2. Display more detailed results:

   `metrocluster check run`

   `metrocluster check aggregate show`

   `metrocluster check cluster show`

   `metrocluster check config-replication show`

   `metrocluster check lif show`

   `metrocluster check node show`

   The `metrocluster check show` commands show the results of the most recent `metrocluster check run` command. You should always run the `metrocluster check run` command prior to using the `metrocluster check show` commands so that the information displayed is current.

   The following example shows the `metrocluster check aggregate show` command output for a healthy four-node MetroCluster configuration:

```
cluster_A::> metrocluster check aggregate show

Last Checked On: 8/5/2014 00:42:58

Node                  Aggregate                  Check
Result
---------------       --------------------       ---------------------
---------
controller_A_1        controller_A_1_aggr0
                                                 mirroring-status
ok
                                                 disk-pool-allocation
ok
```

```
                                                      ownership-state
ok
                      controller_A_1_aggr1
                                                      mirroring-status
ok
                                                      disk-pool-allocation
ok
                                                      ownership-state
ok
                      controller_A_1_aggr2
                                                      mirroring-status
ok
                                                      disk-pool-allocation
ok
                                                      ownership-state
ok


controller_A_2        controller_A_2_aggr0
                                                      mirroring-status
ok
                                                      disk-pool-allocation
ok
                                                      ownership-state
ok
                      controller_A_2_aggr1
                                                      mirroring-status
ok
                                                      disk-pool-allocation
ok
                                                      ownership-state
ok
                      controller_A_2_aggr2
                                                      mirroring-status
ok
                                                      disk-pool-allocation
ok
                                                      ownership-state
ok

18 entries were displayed.
```

The following example shows the `metrocluster check cluster show` command output for a healthy four-node MetroCluster configuration. It indicates that the clusters are ready to perform a negotiated switchover if necessary.

```
Last Checked On: 9/13/2017 20:47:04

Cluster               Check                            Result
--------------------- ------------------------------- ---------
mccint-fas9000-0102
                      negotiated-switchover-ready     not-applicable
                      switchback-ready                not-applicable
                      job-schedules                   ok
                      licenses                        ok
                      periodic-check-enabled          ok
mccint-fas9000-0304
                      negotiated-switchover-ready     not-applicable
                      switchback-ready                not-applicable
                      job-schedules                   ok
                      licenses                        ok
                      periodic-check-enabled          ok
10 entries were displayed.
```

Related information

Disk and aggregate management

Network and LIF management

Checking for MetroCluster configuration errors with Config Advisor

You can go to the NetApp Support Site and download the Config Advisor tool to check for common configuration errors.

Config Advisor is a configuration validation and health check tool. You can deploy it at both secure sites and non-secure sites for data collection and system analysis.

Support for Config Advisor is limited, and available only online.

1. Go to the Config Advisor download page and download the tool.

   NetApp Downloads: Config Advisor

2. Run Config Advisor, review the tool's output and follow the recommendations in the output to address any issues discovered.

Verifying switchover, healing, and switchback

You should verify the switchover, healing, and switchback operations of the MetroCluster configuration.

1. Use the procedures for negotiated switchover, healing, and switchback that are mentioned in the Perform switchover, healing, and switchback.

Protecting configuration backup files

You can provide additional protection for the cluster configuration backup files by specifying a remote URL (either HTTP or FTP) where the configuration backup files will be uploaded in addition to the default locations in the local cluster.

1. Set the URL of the remote destination for the configuration backup files:

   ```
   system configuration backup settings modify URL-of-destination
   ```

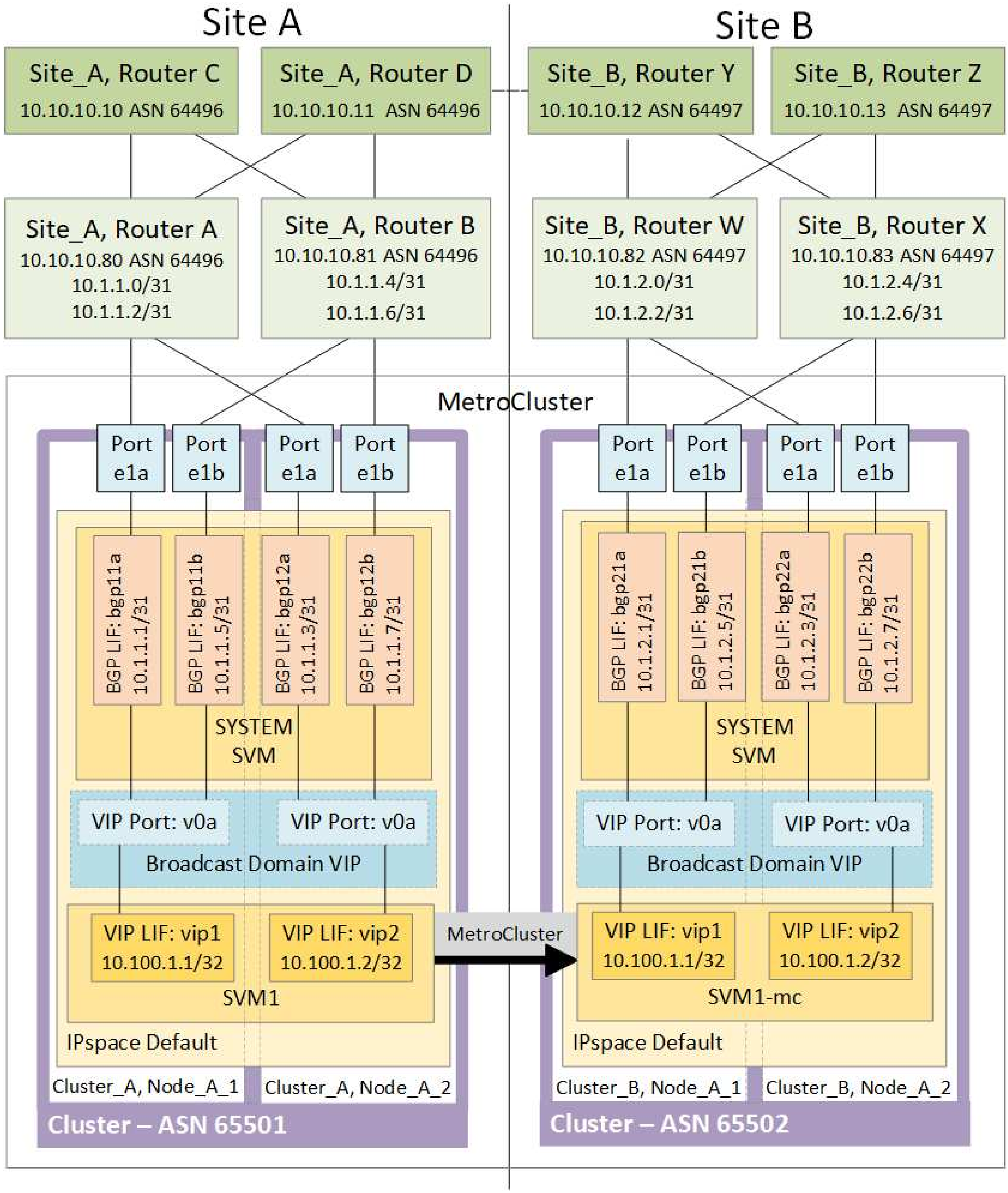
   The Cluster Management with the CLI contains additional information under the section *Managing configuration backups*.

Considerations for using virtual IP and Border Gateway Protocol with a MetroCluster configuration

Beginning with ONTAP 9.5, ONTAP supports layer 3 connectivity using virtual IP (VIP) and Border Gateway Protocol (BGP). The combination VIP and BGP for redundancy in the front-end networking with the back-end MetroCluster redundancy provides a layer 3 disaster recovery solution.

Review the following guidelines and illustration when planning your layer 3 solution. For details on implementing VIP and BGP in ONTAP, refer to the following section:

Configuring virtual IP (VIP) LIFs



ONTAP limitations

ONTAP does not automatically verify that all nodes on both sites of the MetroCluster configuration are configured with BGP peering.

ONTAP does not perform route aggregation but announces all individual virtual LIF IPs as unique host routes at all times.

ONTAP does not support true AnyCast — only a single node in the cluster presents a specific virtual LIF IP (but is accepted by all physical interfaces, regardless of whether they are BGP LIFs, provided the physical port is part of the correct IPspace). Different LIFs can migrate independently of each other to different hosting nodes.

Guidelines for using this Layer 3 solution with a MetroCluster configuration

You must configure your BGP and VIP correctly to provide the required redundancy.

Simpler deployment scenarios are preferred over more complex architectures (for example, a BGP peering router is reachable across an intermediate, non-BGP router). However, ONTAP does not enforce network design or topology restrictions.

VIP LIFs only cover the frontend/data network.

Depending on your version of ONTAP, you must configure BGP peering LIFs in the node SVM, not the system or data SVM. In ONTAP 9.8, the BGP LIFs are visible in the cluster (system) SVM and the node SVMs are no longer present.

Each data SVM requires the configuration of all potential first hop gateway addresses (typically, the BGP router peering IP address), so that the return data path is available if a LIF migration or MetroCluster failover occurs.

BGP LIFs are node specific, similar to intercluster LIFs — each node has a unique configuration, which does not need to be replicated to DR site nodes.

The existence of the v0a (v0b and so on.) continuously validates the connectivity, guaranteeing that a LIF migrate or failover succeeds (unlike L2, where a broken configuration is only visible after the outage).

A major architectural difference is that clients should no longer share the same IP subnet as the VIP of data SVMs. An L3 router with appropriate enterprise grade resiliency and redundancy features enabled (for example, VRRP/HSRP) should be on the path between storage and clients for the VIP to operate correctly.

The reliable update process of BGP allows for smoother LIF migrations because they are marginally faster and have a lower chance of interruption to some clients.

You can configure BGP to detect some classes of network or switch misbehaviors faster than LACP, if configured accordingly.

External BGP (EBGP) uses different AS numbers between ONTAP node(s) and peering routers and is the preferred deployment to ease route aggregation and redistribution on the routers. Internal BGP (IBGP) and the use of route reflectors is not impossible but outside the scope of a straightforward VIP setup.

After deployment, you must check that the data SVM is accessible when the associated virtual LIF is migrated between all nodes on each site (including MetroCluster switchover) to verify the correct configuration of the static routes to the same data SVM.

VIP works for most IP-based protocols (NFS, SMB, iSCSI).

Testing the MetroCluster configuration

You can test failure scenarios to confirm the correct operation of the MetroCluster configuration.

Verifying negotiated switchover

You can test a negotiated (planned) switchover operation to confirm uninterrupted data availability.

This test validates that data availability is not affected (except for SMB and Fibre Channel protocols) by switching the cluster over to the second data center.

This test should take about 30 minutes.

This procedure has the following expected results:

- The `metrocluster switchover` command will present a warning prompt.

  If you respond **`yes`** to the prompt, the site the command is issued from will switch over the partner site.

For MetroCluster IP configurations:

- For ONTAP 9.4 and earlier:
  - Mirrored aggregates will become degraded after the negotiated switchover.
- For ONTAP 9.5 and later:
  - Mirrored aggregates will remain in normal state if the remote storage is accessible.
  - Mirrored aggregates will become degraded after the negotiated switchover if access to the remote storage is lost.
- For ONTAP 9.8 and later:
  - Unmirrored aggregates that are located at the disaster site will become unavailable if access to the remote storage is lost. This might lead to a controller outage.

Steps

1. Confirm that all nodes are in the configured state and normal mode:

   `metrocluster node show`

   ```
   cluster_A::>  metrocluster node show

   Cluster                        Configuration State    Mode
   ------------------------------ ----------------------
   ------------------------
    Local: cluster_A               configured             normal
   Remote: cluster_B               configured             normal
   ```

2. Begin the switchover operation:

   `metrocluster switchover`

```
cluster_A::> metrocluster switchover
Warning: negotiated switchover is about to start. It will stop all the
data Vservers on cluster "cluster_B" and
automatically re-start them on cluster "cluster_A". It will finally
gracefully shutdown cluster "cluster_B".
```

3. Confirm that the local cluster is in the configured state and switchover mode:

   `metrocluster node show`

```
cluster_A::>  metrocluster node show

Cluster                        Configuration State    Mode
------------------------------ ----------------------
------------------------
Local: cluster_A                configured             switchover
Remote: cluster_B               not-reachable          -
              configured             normal
```

4. Confirm that the switchover operation was successful:

   `metrocluster operation show`

```
cluster_A::> metrocluster operation show
  Operation: switchover
      State: successful
 Start Time: 2/6/2016 13:28:50
   End Time: 2/6/2016 13:29:41
     Errors: -
```

5. Use the `vserver show` and `network interface show` commands to verify that DR SVMs and LIFs have come online.

Verifying healing and manual switchback

You can test the healing and manual switchback operations to verify that data availability is not affected (except for SMB and Solaris FC configurations) by switching back the cluster to the original data center after a negotiated switchover.

This test should take about 30 minutes.

The expected result of this procedure is that services should be switched back to their home nodes.

Steps

1. Verify that healing is completed:

```
metrocluster node show
```

The following example shows the successful completion of the command:

```
cluster_A::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node               State          Mirroring Mode
----- ------- ------------------ -------------- ---------
--------------------
1     cluster_A
              node_A_1         configured     enabled   heal roots
completed
      cluster_B
              node_B_2         unreachable    -         switched over
42 entries were displayed.
```

2. Verify that all aggregates are `mirrored`:

```
storage aggregate show
```

The following example shows that all aggregates have a RAID Status of mirrored:

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate Size     Available Used% State   #Vols  Nodes       RAID
Status
--------- -------- --------- ----- ------- ------ -----------
------------
data_cluster
            4.19TB    4.13TB    2% online       8 node_A_1    raid_dp,
                                                              mirrored,
                                                              normal
root_cluster
           715.5GB   212.7GB   70% online       1 node_A_1    raid4,
                                                              mirrored,
                                                              normal
cluster_B Switched Over Aggregates:
Aggregate Size     Available Used% State   #Vols  Nodes       RAID
Status
--------- -------- --------- ----- ------- ------ -----------
------------
data_cluster_B
            4.19TB    4.11TB    2% online       5 node_A_1    raid_dp,
                                                              mirrored,
                                                              normal
root_cluster_B    -         -     - unknown      - node_A_1   -
```

3. Boot nodes from the disaster site.
4. Check the status of switchback recovery:

   `metrocluster node show`

```
cluster_A::> metrocluster node show
DR                               Configuration  DR
Group Cluster Node               State          Mirroring Mode
----- ------- ------------------ -------------- ---------
--------------------
1     cluster_A
             node_A_1            configured     enabled   heal roots
completed
      cluster_B
             node_B_2            configured     enabled   waiting for
switchback
                                                          recovery
2 entries were displayed.
```

5. Perform the switchback:

   `metrocluster switchback`

```
cluster_A::> metrocluster switchback
[Job 938] Job succeeded: Switchback is successful.Verify switchback
```

6. Confirm status of the nodes:

   `metrocluster node show`

```
cluster_A::> metrocluster node show
DR                                Configuration  DR
Group Cluster Node                State          Mirroring Mode
----- ------- ------------------ -------------- ---------
--------------------
1     cluster_A
              node_A_1         configured     enabled   normal
      cluster_B
              node_B_2         configured     enabled   normal

2 entries were displayed.
```

7. Confirm the status:

   `metrocluster operation show`

   The output should show a successful state.

```
cluster_A::> metrocluster operation show
  Operation: switchback
      State: successful
 Start Time: 2/6/2016 13:54:25
   End Time: 2/6/2016 13:56:15
     Errors: -
```

Loss of a single FC-to-SAS bridge

You can test the failure of a single FC-to-SAS bridge to make sure there is no single point of failure.

This test should take about 15 minutes.

This procedure has the following expected results:

- Errors should be generated as the bridge is switched off.

- No failover or loss of service should occur.
- Only one path from the controller module to the drives behind the bridge is available.

> Beginning with ONTAP 9.8, the `storage bridge` command is replaced with `system bridge`. The following steps show the `storage bridge` command, but if you are running ONTAP 9.8 or later, the `system bridge` command is preferred.

Steps

1. Turn off the power supplies of the bridge.
2. Confirm that the bridge monitoring indicates an error:

   `storage bridge show`

   ```
   cluster_A::> storage bridge show

                                                               Is
   Monitor
   Bridge     Symbolic Name Vendor  Model     Bridge WWN       Monitored
   Status
   ---------- ------------- ------- --------- ---------------- ---------
   -------
   ATTO_10.65.57.145
            bridge_A_1    Atto    FibreBridge 6500N
                                              200000108662d46c true
   error
   ```

3. Confirm that drives behind the bridge are available with a single path:

   `storage disk error show`

```
cluster_A::> storage disk error show
Disk             Error Type        Error Text
---------------- -----------------
--------------------------------------------
1.0.0            onedomain         1.0.0 (5000cca057729118): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
1.0.1            onedomain         1.0.1 (5000cca057727364): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
1.0.2            onedomain         1.0.2 (5000cca05772e9d4): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
...
1.0.23           onedomain         1.0.23 (5000cca05772e9d4): All paths
to this array LUN are connected to the same fault domain. This is a
single point of failure.
```

Verifying operation after power line disruption

You can test the MetroCluster configuration's response to the failure of a PDU.

The best practice is for each power supply unit (PSU) in a component to be connected to a separate power supply. If both PSUs are connected to the same power distribution unit (PDU) and an electrical disruption occurs, the site could down and a complete shelf might become unavailable. Failure of one power line is tested to confirm that there is no cabling mismatch that could cause a service disruption.

This test should take about 15 minutes.

This test requires turning off power to all left-hand PDUs and then all right-hand PDUs on all of the racks containing the MetroCluster components.

This procedure has the following expected results:

- Errors should be generated as the PDUs are disconnected.
- No failover or loss of service should occur.

Steps

1. Turn off the power of the PDUs on the left-hand side of the rack containing the MetroCluster components.
2. Monitor the result on the console by using the `system environment sensors show -state fault` and `storage shelf show -errors` commands.

```
cluster_A::> system environment sensors show -state fault

Node Sensor              State Value/Units Crit-Low Warn-Low Warn-Hi
Crit-Hi
---- --------------------- ------ ----------- -------- -------- -------
-------
node_A_1
        PSU1             fault
                             PSU_OFF
        PSU1 Pwr In OK  fault
                             FAULT
node_A_2
        PSU1             fault
                             PSU_OFF
        PSU1 Pwr In OK  fault
                             FAULT
4 entries were displayed.

cluster_A::> storage shelf show -errors
    Shelf Name: 1.1
     Shelf UID: 50:0a:09:80:03:6c:44:d5
 Serial Number: SHFHU1443000059

Error Type          Description
------------------  ---------------------------
Power               Critical condition is detected in storage shelf
power supply unit "1". The unit might fail.Reconnect PSU1
```

3. Turn the power back on to the left-hand PDUs.
4. Make sure that ONTAP clears the error condition.
5. Repeat the previous steps with the right-hand PDUs.

Verifying operation after loss of a single storage shelf

You can test the failure of a single storage shelf to verify that there is no single point of failure.

This procedure has the following expected results:

- An error message should be reported by the monitoring software.
- No failover or loss of service should occur.
- Mirror resynchronization starts automatically after the hardware failure is restored.

Steps

1. Check the storage failover status:

   `storage failover show`

```
cluster_A::> storage failover show

Node           Partner        Possible State Description
-------------- -------------- --------
-------------------------------------
node_A_1       node_A_2       true     Connected to node_A_2
node_A_2       node_A_1       true     Connected to node_A_1
2 entries were displayed.
```

2. Check the aggregate status:

```
storage aggregate show
```

```
cluster_A::> storage aggregate show

cluster Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
node_A_1data01_mirrored
            4.15TB    3.40TB   18% online       3 node_A_1
raid_dp,

mirrored,

normal
node_A_1root
           707.7GB   34.29GB   95% online       1 node_A_1
raid_dp,

mirrored,

normal
node_A_2_data01_mirrored
            4.15TB    4.12TB    1% online       2 node_A_2
raid_dp,

mirrored,

normal
node_A_2_data02_unmirrored
            2.18TB    2.18TB    0% online       1 node_A_2
raid_dp,

normal
node_A_2_root
           707.7GB   34.27GB   95% online       1 node_A_2
raid_dp,

mirrored,

normal
```

3. Verify that all data SVMs and data volumes are online and serving data:

```
vserver show -type data
```

```
network interface show -fields is-home false
```

```
volume show !vol0,!MDV*
```

```
cluster_A::> vserver show -type data

cluster_A::> vserver show -type data
                               Admin      Operational Root
Vserver     Type    Subtype    State      State       Volume
Aggregate
----------- ------- ---------- ---------- ----------- ----------
----------
SVM1        data    sync-source           running     SVM1_root
node_A_1_data01_mirrored
SVM2        data    sync-source           running     SVM2_root
node_A_2_data01_mirrored

cluster_A::> network interface show -fields is-home false
There are no entries matching your query.

cluster_A::> volume show !vol0,!MDV*
Vserver   Volume       Aggregate    State      Type       Size
Available Used%
--------- ------------ ------------ ---------- ---- ----------
---------- -----
SVM1
          SVM1_root
                       node_A_1data01_mirrored
                                    online     RW         10GB
9.50GB    5%
SVM1
          SVM1_data_vol
                       node_A_1data01_mirrored
                                    online     RW         10GB
9.49GB    5%
SVM2
          SVM2_root
                       node_A_2_data01_mirrored
                                    online     RW         10GB
9.49GB    5%
SVM2
          SVM2_data_vol
                       node_A_2_data02_unmirrored
                                    online     RW          1GB
972.6MB    5%
```

4. Identify a shelf in Pool 1 for node node_A_2 to power off to simulate a sudden hardware failure:

   `storage aggregate show -r -node *node-name* !*root`

   The shelf you select must contain drives that are part of a mirrored data aggregate.

   In the following example, shelf ID 31 is selected to fail.

```
cluster_A::> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
 Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirrored) (block
checksums)
  Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
   RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- --------------------------- ---- ----- ------ --------
-------- ----------
     dparity  2.30.3                       0   BSAS    7200  827.7GB
828.0GB (normal)
     parity   2.30.4                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.6                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.8                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.5                       0   BSAS    7200  827.7GB
828.0GB (normal)

  Plex: /node_A_2_data01_mirrored/plex4 (online, normal, active, pool1)
   RAID Group /node_A_2_data01_mirrored/plex4/rg0 (normal, block
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- --------------------------- ---- ----- ------ --------
-------- ----------
     dparity  1.31.7                       1   BSAS    7200  827.7GB
828.0GB (normal)
     parity   1.31.6                       1   BSAS    7200  827.7GB
828.0GB (normal)
     data     1.31.3                       1   BSAS    7200  827.7GB
828.0GB (normal)
```

```
     data     1.31.4                        1   BSAS    7200  827.7GB
828.0GB (normal)
     data     1.31.5                        1   BSAS    7200  827.7GB
828.0GB (normal)

 Aggregate: node_A_2_data02_unmirrored (online, raid_dp) (block
checksums)
  Plex: /node_A_2_data02_unmirrored/plex0 (online, normal, active,
pool0)
   RAID Group /node_A_2_data02_unmirrored/plex0/rg0 (normal, block
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- --------------------------- ---- ----- ------ --------
-------- ----------
     dparity  2.30.12                       0   BSAS    7200  827.7GB
828.0GB (normal)
     parity   2.30.22                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.21                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.20                       0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.14                       0   BSAS    7200  827.7GB
828.0GB (normal)
15 entries were displayed.
```

5. Physically power off the shelf that you selected.
6. Check the aggregate status again:

   `storage aggregate`

   `storage aggregate show -r -node node_A_2 !*root`

   The aggregate with drives on the powered-off shelf should have a “degraded” RAID status, and drives on the affected plex should have a “failed” status, as shown in the following example:

```
cluster_A::> storage aggregate show
Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
node_A_1data01_mirrored
            4.15TB    3.40TB   18% online       3 node_A_1
```

```
raid_dp,

mirrored,

normal
node_A_1root
           707.7GB   34.29GB   95% online       1 node_A_1
raid_dp,

mirrored,

normal
node_A_2_data01_mirrored
            4.15TB    4.12TB    1% online       2 node_A_2
raid_dp,

mirror

degraded
node_A_2_data02_unmirrored
            2.18TB    2.18TB    0% online       1 node_A_2
raid_dp,

normal
node_A_2_root
           707.7GB   34.27GB   95% online       1 node_A_2
raid_dp,

mirror

degraded
cluster_A::> storage aggregate show -r -node node_A_2 !*root
Owner Node: node_A_2
 Aggregate: node_A_2_data01_mirrored (online, raid_dp, mirror degraded)
(block checksums)
  Plex: /node_A_2_data01_mirrored/plex0 (online, normal, active, pool0)
   RAID Group /node_A_2_data01_mirrored/plex0/rg0 (normal, block
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- --------------------------- ---- ----- ------ --------
-------- ----------
     dparity  2.30.3                       0   BSAS    7200  827.7GB
828.0GB (normal)
```

```
     parity   2.30.4                           0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.6                           0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.8                           0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.5                           0   BSAS    7200  827.7GB
828.0GB (normal)

  Plex: /node_A_2_data01_mirrored/plex4 (offline, failed, inactive,
pool1)
   RAID Group /node_A_2_data01_mirrored/plex4/rg0 (partial, none
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- -------------------------- ---- ----- ------ --------
-------- ----------
     dparity  FAILED                       -   -          -  827.7GB
- (failed)
     parity   FAILED                       -   -          -  827.7GB
- (failed)
     data     FAILED                       -   -          -  827.7GB
- (failed)
     data     FAILED                       -   -          -  827.7GB
- (failed)
     data     FAILED                       -   -          -  827.7GB
- (failed)

 Aggregate: node_A_2_data02_unmirrored (online, raid_dp) (block
checksums)
  Plex: /node_A_2_data02_unmirrored/plex0 (online, normal, active,
pool0)
   RAID Group /node_A_2_data02_unmirrored/plex0/rg0 (normal, block
checksums)
                                                              Usable
Physical
     Position Disk                        Pool Type     RPM     Size
Size Status
     -------- -------------------------- ---- ----- ------ --------
-------- ----------
     dparity  2.30.12                      0   BSAS    7200  827.7GB
828.0GB (normal)
     parity   2.30.22                      0   BSAS    7200  827.7GB
828.0GB (normal)
```

```
     data     2.30.21                      0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.20                      0   BSAS    7200  827.7GB
828.0GB (normal)
     data     2.30.14                      0   BSAS    7200  827.7GB
828.0GB (normal)
15 entries were displayed.
```

7. Verify that the data is being served and that all volumes are still online:

```
vserver show -type data
```

```
network interface show -fields is-home false
```

```
volume show !vol0,!MDV*
```

```
cluster_A::> vserver show -type data

cluster_A::> vserver show -type data
                               Admin      Operational Root
Vserver     Type    Subtype    State      State       Volume
Aggregate
----------- ------- ---------- ---------- ----------- ----------
----------
SVM1        data    sync-source           running     SVM1_root
node_A_1_data01_mirrored
SVM2        data    sync-source           running     SVM2_root
node_A_1_data01_mirrored

cluster_A::> network interface show -fields is-home false
There are no entries matching your query.

cluster_A::> volume show !vol0,!MDV*
Vserver   Volume       Aggregate    State      Type       Size
Available Used%
--------- ------------ ------------ ---------- ---- ----------
---------- -----
SVM1
          SVM1_root
                       node_A_1data01_mirrored
                                    online     RW         10GB
9.50GB    5%
SVM1
          SVM1_data_vol
                       node_A_1data01_mirrored
                                    online     RW         10GB
9.49GB    5%
SVM2
          SVM2_root
                       node_A_1data01_mirrored
                                    online     RW         10GB
9.49GB    5%
SVM2
          SVM2_data_vol
                       node_A_2_data02_unmirrored
                                    online     RW          1GB
972.6MB    5%
```

8. Physically power on the shelf.

   Resynchronization starts automatically.

9. Verify that resynchronization has started:

   `storage aggregate show`

   The affected aggregate should have a “resyncing” RAID status, as shown in the following example:

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate     Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
node_A_1_data01_mirrored
            4.15TB    3.40TB   18% online       3 node_A_1
raid_dp,

mirrored,

normal
node_A_1_root
           707.7GB   34.29GB   95% online       1 node_A_1
raid_dp,

mirrored,

normal
node_A_2_data01_mirrored
            4.15TB    4.12TB    1% online       2 node_A_2
raid_dp,

resyncing
node_A_2_data02_unmirrored
            2.18TB    2.18TB    0% online       1 node_A_2
raid_dp,

normal
node_A_2_root
           707.7GB   34.27GB   95% online       1 node_A_2
raid_dp,

resyncing
```

10. Monitor the aggregate to confirm that resynchronization is complete:

    `storage aggregate show`

    The affected aggregate should have a “normal” RAID status, as shown in the following example:

```
cluster_A::> storage aggregate show
cluster Aggregates:
Aggregate      Size Available Used% State   #Vols  Nodes            RAID
Status
--------- -------- --------- ----- ------- ------ ----------------
------------
node_A_1data01_mirrored
            4.15TB    3.40TB   18% online       3 node_A_1
raid_dp,

mirrored,

normal
node_A_1root
           707.7GB   34.29GB   95% online       1 node_A_1
raid_dp,

mirrored,

normal
node_A_2_data01_mirrored
            4.15TB    4.12TB    1% online       2 node_A_2
raid_dp,

normal
node_A_2_data02_unmirrored
            2.18TB    2.18TB    0% online       1 node_A_2
raid_dp,

normal
node_A_2_root
           707.7GB   34.27GB   95% online       1 node_A_2
raid_dp,

resyncing
```

Remove MetroCluster configurations

If you need to remove the MetroCluster configuration, contact technical support.

Contact NetApp technical support and reference the appropriate guide for your configuration from How to remove nodes from a MetroCluster configuration - Resolution Guide.

You cannot reverse the MetroCluster unconfiguration. This process should only be done with the assistance of technical support. After removing the MetroCluster configuration, all disk connectivity and interconnects should be adjusted to be in a supported state.

How to use the Active IQ Unified Manager and ONTAP System Manager for further configuration and monitoring

Use the Active IQ Unified Manager and ONTAP System Manager for further configuration and monitoring

The Active IQ Unified Manager and ONTAP System Manager can be used for GUI management of the clusters and monitoring the configuration.

Each node has ONTAP System Manager pre-installed. To load System Manager, enter the cluster management LIF address as the URL in a web browser that has connectivity to the node.

You can also use Active IQ Unified Manager to monitor the MetroCluster configuration.

Related information

Active IQ Unified Manager Documentation

Synchronize the system time using NTP

Each cluster needs its own Network Time Protocol (NTP) server to synchronize the time between the nodes and their clients.

About this task

- You cannot modify the time zone settings for a failed node or the partner node after takeover occurs.
- Each cluster in the stretch MetroCluster configuration should have its own separate NTP server or servers used by the nodes at that MetroCluster site.
- If you are using the MetroCluster Tiebreaker software, it should also have its own separate NTP server.

Depending on your ONTAP version, you can configure the NTP from the **Cluster** or **Insights** tab in the System Manager UI.

Cluster

In System Manager, you can configure the NTP from the **Cluster** tab using two different options, depending on your ONTAP version:

ONTAP 9.8 or later:

Use the following steps to synchronize the NTP from the **Cluster** tab in ONTAP 9.8 or later.

Steps

1. Go to **Cluster > Overview**
2. Then select the 

option and select **Edit**.
3. In the **Edit Cluster Details** window, select the **+Add** option below NTP Servers.
4. Add the name, location, and specify the IP address of the time server.
5. Then, select **Save**.
6. Repeat the steps for any additional time servers.

ONTAP 9.11.1 or later:

Use the following steps to synchronize the NTP from the **Insights** window in the **Cluster** tab in ONTAP 9.11.1 or later.

Steps

1. Go to **Cluster > Overview**
2. Scroll down to the **Insights** window on the page, locate **Too few NTP servers are configured**, and then select **Fix It**.
3. Specify the IP address of the time server, and then select **Save**.
4. Repeat the previous step for any additional time servers.

Insights

In ONTAP 9.11.1 or later, you can also configure the NTP by using the **Insights** tab in System Manager:

Steps

1. Go to the **Insights** tab in the System Manager UI.
2. Scroll down to **Too few NTP servers are configured** and select **Fix It**.
3. Specify the IP address of the time server, and then select **Save**.
4. Repeat the previous step for any additional time servers.

Considerations when using ONTAP in a MetroCluster configuration

When using ONTAP in a MetroCluster configuration, you should be aware of certain considerations for licensing, peering to clusters outside the MetroCluster configuration, performing volume operations, NVFAIL operations, and other ONTAP operations.

Licensing considerations

- Both sites should be licensed for the same site-licensed features.
- All nodes should be licensed for the same node-locked features.

SnapMirror consideration

- SnapMirror SVM disaster recovery is only supported on MetroCluster configurations running versions of ONTAP 9.5 or later.

FlexCache support in a MetroCluster configuration

Beginning with ONTAP 9.7, FlexCache volumes are supported on MetroCluster configurations. You should be aware of requirements for manual repeering after switchover or switchback operations.

SVM repeering after switchover when FlexCache origin and cache are within the same MetroCluster site

After a negotiated or unplanned switchover, any SVM FlexCache peering relationship within the cluster must be manually configured.

For example, SVMs vs1 (cache) and vs2 (origin) are on site_A. These SVMs are peered.

After switchover, SVMs vs1-mc and vs2-mc are activated at the partner site (site_B). They must be manually repeered for FlexCache to work using the `vserver peer repeer` command.

SVM repeering after switchover or switchback when a FlexCache destination is on a third cluster and in disconnected mode

For FlexCache relationships to a cluster outside of the MetroCluster configuration, the peering must always be manually reconfigured after a switchover when the clusters involved are in a disconnected mode during switchover.

For example:

- One end of the FlexCache (cache_1 on vs1) resides on MetroCluster site_A has one end of the FlexCache
- The other end of the FlexCache (origin_1 on vs2) resides on site_C (not in the MetroCluster configuration)

When switchover is triggered, and if site_A and site_C are not connected, you must manually repeer the SVMs on site_B (the switchover cluster) and site_C using the `vserver peer repeer` command after the switchover.

When switchback is performed, you must again repeer the SVMs on site_A (the original cluster) and site_C.

FabricPool support in MetroCluster configurations

Beginning with ONTAP 9.7, MetroCluster configurations support FabricPool storage tiers.

For general information on using FabricPools, see the Disks and aggregates management.

Considerations when using FabricPools

- The clusters must have FabricPool licenses with matching capacity limits.

- The clusters must have IPspaces with matching names.

  This can be the default IPspace, or an IP space an administer has created. This IPspace will be used for FabricPool object store configuration setups.

- For the selected IPspace, each cluster must have an intercluster LIF defined that can reach the external object store.

Configuring an aggregate for use in a mirrored FabricPool

> Before you configure the aggregate you must set up object stores as described in "Setting up object stores for FabricPool in a MetroCluster configuration" in the Disks and aggregates management.

To configure an aggregate for use in a FabricPool:

1. Create the aggregate or select an existing aggregate.
2. Mirror the aggregate as a typical mirrored aggregate within the MetroCluster configuration.
3. Create the FabricPool mirror with the aggregate, as described in the Disks and aggregates management:
   a. Attach a primary object store.

      This object store is physically closer to the cluster.

   b. Add a mirror object store.

      This object store is physically further away from the cluster than the primary object store.

FlexGroup support in MetroCluster configurations

Beginning with ONTAP 9.6 MetroCluster configurations support FlexGroup volumes.

Job schedules in a MetroCluster configuration

In ONTAP 9.3 and later, user-created job schedules are automatically replicated between clusters in a MetroCluster configuration. If you create, modify, or delete a job schedule on a cluster, the same schedule is automatically created on the partner cluster, using Configuration Replication Service (CRS).

> System-created schedules are not replicated and you must manually perform the same operation on the partner cluster so that job schedules on both clusters are identical.

Cluster peering from the MetroCluster site to a third cluster

Because the peering configuration is not replicated, if you peer one of the clusters in the MetroCluster configuration to a third cluster outside of that configuration, you must also configure the peering on the partner MetroCluster cluster. This is so that peering can be maintained if a switchover occurs.

The non-MetroCluster cluster must be running ONTAP 8.3 or later. If not, peering is lost if a switchover occurs even if the peering has been configured on both MetroCluster partners.

LDAP client configuration replication in a MetroCluster configuration

An LDAP client configuration created on a storage virtual machine (SVM) on a local cluster is replicated to its partner data SVM on the remote cluster. For example, if the LDAP client configuration is created on the admin SVM on the local cluster, then it is replicated to all the admin data SVMs on the remote cluster. This MetroCluster feature is intentional so that the LDAP client configuration is active on all the partner SVMs on the remote cluster.

Networking and LIF creation guidelines for MetroCluster configurations

You should be aware of how LIFs are created and replicated in a MetroCluster configuration. You must also know about the requirement for consistency so that you can make proper decisions when configuring your network.

Related information

ONTAP concepts

IPspace object replication and subnet configuration requirements

You should be aware of the requirements for replicating IPspace objects to the partner cluster and for configuring subnets and IPv6 in a MetroCluster configuration.

IPspace replication

You must consider the following guidelines while replicating IPspace objects to the partner cluster:

- The IPspace names of the two sites must match.
- IPspace objects must be manually replicated to the partner cluster.

  Any storage virtual machines (SVMs) that are created and assigned to an IPspace before the IPspace is replicated will not be replicated to the partner cluster.

Subnet configuration

You must consider the following guidelines while configuring subnets in a MetroCluster configuration:

- Both clusters of the MetroCluster configuration must have a subnet in the same IPspace with the same subnet name, subnet, broadcast domain, and gateway.
- The IP ranges of the two clusters must be different.

  In the following example, the IP ranges are different:

```
cluster_A::> network subnet show

IPspace: Default
Subnet                    Broadcast                    Avail/
Name      Subnet          Domain    Gateway            Total    Ranges
--------- ---------------- --------- ------------      -------
---------------
subnet1   192.168.2.0/24   Default   192.168.2.1       10/10
192.168.2.11-192.168.2.20

cluster_B::> network subnet show
 IPspace: Default
Subnet                    Broadcast                    Avail/
Name      Subnet          Domain    Gateway            Total    Ranges
--------- ---------------- --------- ------------     --------
---------------
subnet1   192.168.2.0/24   Default   192.168.2.1       10/10
192.168.2.21-192.168.2.30
```

IPv6 configuration

If IPv6 is configured on one site, IPv6 must be configured on the other site as well.

Requirements for LIF creation in a MetroCluster configuration

You should be aware of the requirements for creating LIFs when configuring your network in a MetroCluster configuration.

You must consider the following guidelines when creating LIFs:

- Fibre Channel: You must use stretched VSAN or stretched fabrics.
- IP/iSCSI: You must use layer 2 stretched network.
- ARP broadcasts: You must enable ARP broadcasts between the two clusters.
- Duplicate LIFs: You must not create multiple LIFs with the same IP address (duplicate LIFs) in an IPspace.
- NFS and SAN configurations: You must use different storage virtual machines (SVMs) for both the unmirrored and mirrored aggregates.
- You should create a subnet object before you create a LIF. A subnet object enables ONTAP to determine failover targets on the destination cluster because it has an associated broadcast domain.

Verify LIF creation

You can confirm the successful creation of a LIF in a MetroCluster configuration by running the `metrocluster check lif show` command. If you encounter any issues while creating the LIF, you can use the `metrocluster check lif repair-placement` command to fix the issues.

LIF replication and placement requirements and issues

You should be aware of the LIF replication requirements in a MetroCluster configuration. You should also know how a replicated LIF is placed on a partner cluster, and you should be aware of the issues that occur when LIF replication or LIF placement fails.

Replication of LIFs to the partner cluster

When you create a LIF on a cluster in a MetroCluster configuration, the LIF is replicated on the partner cluster. LIFs are not placed on a one-to-one name basis. For availability of LIFs after a switchover operation, the LIF placement process verifies that the ports are able to host the LIF based on reachability and port attribute checks.

The system must meet the following conditions to place the replicated LIFs on the partner cluster:

Condition	LIF type: FC	LIF type: IP/iSCSI
Node identification	ONTAP attempts to place the replicated LIF on the disaster recovery (DR) partner of the node on which it was created. If the DR partner is unavailable, the DR auxiliary partner is used for placement.	ONTAP attempts to place the replicated LIF on the DR partner of the node on which it was created. If the DR partner is unavailable, the DR auxiliary partner is used for placement.
Port identification	ONTAP identifies the connected FC target ports on the DR cluster.	The ports on the DR cluster that are in the same IPspace as the source LIF are selected for a reachability check. If there are no ports in the DR cluster in the same IPspace, the LIF cannot be placed. All of the ports in the DR cluster that are already hosting a LIF in the same IPspace and subnet are automatically marked as reachable; and can be used for placement. These ports are not included in the reachability check.

Reachability check	Reachability is determined by checking for the connectivity of the source fabric WWN on the ports in the DR cluster. If the same fabric is not present at the DR site, the LIF is placed on a random port on the DR partner.	Reachability is determined by the response to an Address Resolution Protocol (ARP) broadcast from each previously identified port on the DR cluster to the source IP address of the LIF to be placed. For reachability checks to succeed, ARP broadcasts must be allowed between the two clusters. Each port that receives a response from the source LIF will be marked as possible for placement.
Port selection	ONTAP categorizes the ports based on attributes such as adapter type and speed, and then selects the ports with matching attributes. If no ports with matching attributes are found, the LIF is placed on a random connected port on the DR partner.	From the ports that are marked as reachable during the reachability check, ONTAP prefers ports that are in the broadcast domain that is associated with the subnet of the LIF. If there are no network ports available on the DR cluster that are in the broadcast domain that is associated with the subnet of the LIF, then ONTAP selects ports that have reachability to the source LIF. If there are no ports with reachability to the source LIF, a port is selected from the broadcast domain that is associated with the subnet of the source LIF, and if no such broadcast domain exists, a random port is selected. ONTAP categorizes the ports based on attributes such as adapter type, interface type, and speed, and then selects the ports with matching attributes.
LIF placement	From the reachable ports, ONTAP selects the least loaded port for placement.	From the selected ports, ONTAP selects the least loaded port for placement.

Placement of replicated LIFs when the DR partner node is down

When an iSCSI or FC LIF is created on a node whose DR partner has been taken over, the replicated LIF is placed on the DR auxiliary partner node. After a subsequent giveback operation, the LIFs are not automatically moved to the DR partner. This can lead to LIFs being concentrated on a single node in the partner cluster. During a MetroCluster switchover operation, subsequent attempts to map LUNs belonging to the storage virtual machine (SVM) fail.

You should run the `metrocluster check lif show` command after a takeover operation or giveback operation to verify that the LIF placement is correct. If errors exist, you can run the `metrocluster check lif repair-placement` command to resolve the issues.

LIF placement errors

LIF placement errors that are displayed by the `metrocluster check lif show` command are retained after a switchover operation. If the `network interface modify`, `network interface rename`, or `network interface delete` command is issued for a LIF with a placement error, the error is removed and does not appear in the output of the `metrocluster check lif show` command.

LIF replication failure

You can also check whether LIF replication was successful by using the `metrocluster check lif show` command. An EMS message is displayed if LIF replication fails.

You can correct a replication failure by running the `metrocluster check lif repair-placement` command for any LIF that fails to find a correct port. You should resolve any LIF replication failures as soon as possible to verify the availability of LIF during a MetroCluster switchover operation.

> Even if the source SVM is down, LIF placement might proceed normally if there is a LIF belonging to a different SVM in a port with the same IPspace and network in the destination SVM.

Volume creation on a root aggregate

The system does not allow the creation of new volumes on the root aggregate (an aggregate with an HA policy of CFO) of a node in a MetroCluster configuration.

Because of this restriction, root aggregates cannot be added to an SVM using the `vserver add-aggregates` command.

SVM disaster recovery in a MetroCluster configuration

Beginning with ONTAP 9.5, active storage virtual machines (SVMs) in a MetroCluster configuration can be used as sources with the SnapMirror SVM disaster recovery feature. The destination SVM must be on the third cluster outside of the MetroCluster configuration.

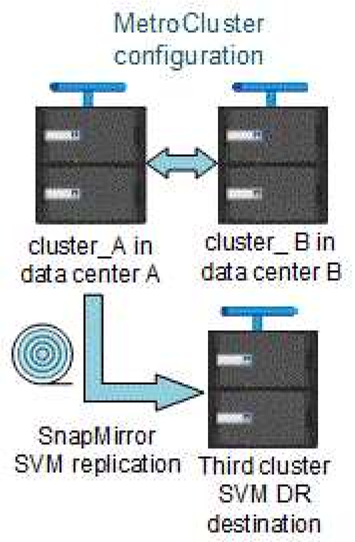
You should be aware of the following requirements and limitations of using SVMs with SnapMirror disaster recovery:

- Only an active SVM within a MetroCluster configuration can be the source of an SVM disaster recovery relationship.

  A source can be a sync-source SVM before switchover or a sync-destination SVM after switchover.

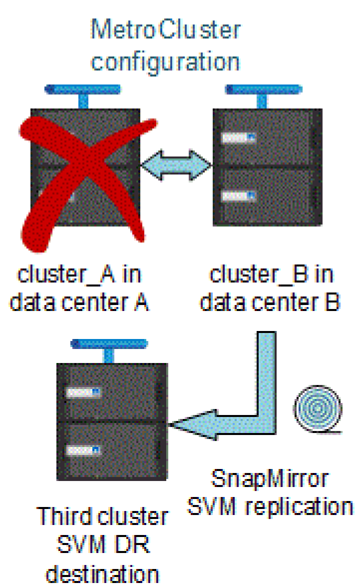
- When a MetroCluster configuration is in a steady state, the MetroCluster sync-destination SVM cannot be the source of an SVM disaster recovery relationship, since the volumes are not online.

  The following image shows the SVM disaster recovery behavior in a steady state:



- When the sync-source SVM is the source of an SVM DR relationship, the source SVM DR relationship information is replicated to the MetroCluster partner.

  This enables the SVM DR updates to continue after a switchover as shown in the following image:



- During the switchover and switchback processes, replication to the SVM DR destination might fail.

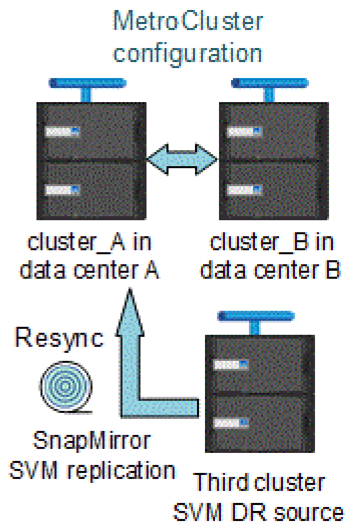
  However, after the switchover or switchback process completes, the next SVM DR scheduled updates will succeed.

See the section “Replicating the SVM configuration” in the Data Protection with the CLI for details on configuring an SVM DR relationship.

SVM resynchronization at a disaster recovery site

During resynchronization, the storage virtual machines (SVMs) disaster recovery (DR) source on the MetroCluster configuration is restored from the destination SVM on the non-MetroCluster site.

During resynchronization, the source SVM (cluster_A) temporarily acts as a destination SVM as shown in the following image:



If an unplanned switchover occurs during resynchronization

Unplanned switchovers that occur during the resynchronization will halt the resynchronization transfer. If an unplanned switchover occurs, the following conditions are true:

- The destination SVM on the MetroCluster site (which was a source SVM prior to resynchronization) remains as a destination SVM. The SVM at the partner cluster will continue to retain its subtype and remain inactive.
- The SnapMirror relationship must be re-created manually with the sync-destination SVM as the destination.
- The SnapMirror relationship does not appear in the SnapMirror show output after a switchover at the survivor site unless a SnapMirror create operation is executed.

Performing switchback after an unplanned switchover during resynchronization

To successfully perform the switchback process, the resynchronization relationship must be broken and deleted. Switchback is not permitted if there are any SnapMirror DR destination SVMs in the MetroCluster configuration or if the cluster has an SVM of subtype “dp-destination”.

Output of the storage disk show and storage shelf show commands in a two-node stretch MetroCluster configuration

In a two-node stretch MetroCluster configuration, the `is-local-attach` field of the `storage disk show` and `storage shelf show` commands shows all of the disks and storage shelves as local, regardless of the node to which they are attached.

Output for the storage aggregate plex show command is indeterminate after a MetroCluster switchover

When you run the `storage aggregate plex show` command after a MetroCluster switchover, the status of plex0 of the switched over root aggregate is indeterminate and is displayed as `failed`. During this time, the switched over root is not updated. The actual status of this plex can only be determined after the MetroCluster healing phase.

Modifying volumes to set the NVFAIL flag in case of switchover

You can modify a volume so that the NVFAIL flag is set on the volume in the event of a MetroCluster switchover. The NVFAIL flag causes the volume to be fenced off from any modification. This is required for volumes that need to be handled as if committed writes to the volume were lost after the switchover.

> In ONTAP versions earlier than 9.0, the NVFAIL flag is used for each switchover. In ONTAP 9.0 and later versions, the unplanned switchover (USO) is used.

Steps

1. Enable MetroCluster configuration to trigger NVFAIL on switchover by setting the `vol -dr-force -nvfail` parameter to "on":

   ```
   vol modify -vserver vserver-name -volume volume-name -dr-force-nvfail on
   ```

Transitioning from a stretch to a fabric-attached MetroCluster configuration

In a fabric-attached MetroCluster configuration, the nodes are in different locations. This geographical difference increases the disaster protection. To transition from a stretch to a fabric-attached MetroCluster configuration, you must add FC switches and, if necessary, FC-to-SAS bridges to the configuration.

- You must disable automatic switchover on both of the clusters by running the `metrocluster modify -auto-switchover-failure-domain auto-disabled` command.
- You must have shut down the nodes.

  This procedure is disruptive.

The MetroCluster configuration must be transitioned on both sites. After upgrading the MetroCluster configuration, you must enable automatic switchover on both the clusters. You also must validate the configuration by running the `metrocluster check run` command.

This procedure gives an overview of the required steps. For detailed steps, you must refer to specific sections in the Fabric-attached MetroCluster installation and configuration. You do not need to do a full installation and configuration.

Steps

1. Prepare for the upgrade by carefully reviewing the "Preparing for the MetroCluster installation" section of the Fabric-attached MetroCluster installation and configuration.
2. Install, cable, and configure the required switches and FC-to-SAS bridges.

   > You should use the procedures in the section "Cabling a fabric-attached MetroCluster configuration" of the Fabric-attached MetroCluster installation and configuration.

3. Refresh the MetroCluster configuration using the following steps.

   Do not use the procedures in the section "Configuring the MetroCluster software in ONTAP" found in the Fabric-attached MetroCluster installation and configuration.

a. Enter advanced privilege mode: + **`set -privilege advanced`**

b. Refresh the MetroCluster configuration: + **`metrocluster configure -refresh true`**

   The following command refreshes the MetroCluster configuration on all the nodes in the DR group that contains controller_A_1:

```
controller_A_1::*> metrocluster configure -refresh true
[Job 009] Job succeeded: Configure is successful.
```

c. Return to admin privilege mode: + **`set -privilege admin`**

4. Check the MetroCluster configuration for errors and verify that it is operational.

   You should use the procedures in the following sections of the Fabric-attached MetroCluster installation and configuration:

   ◦ Checking for MetroCluster configuration errors with Config Advisor
   ◦ Verifying local HA operation
   ◦ Verifying switchover, healing, and switchback

Where to find additional information

You can learn more about the MetroCluster configuration and operation.

MetroCluster and miscellaneous information

Information	Subject
ONTAP 9 Documentation	• All MetroCluster guides
	• A technical overview of the MetroCluster FC configuration and operation. • Best practices for MetroCluster FC configuration.
Fabric-attached MetroCluster installation and configuration	• Fabric-attached MetroCluster architecture • Cabling the configuration • Configuring the FC-to-SAS bridges • Configuring the FC switches • Configuring the MetroCluster in ONTAP
MetroCluster IP installation and configuration: Differences among the ONTAP MetroCluster configurations	• MetroCluster IP architecture • Cabling the configuration • Configuring the MetroCluster in ONTAP

MetroCluster management and disaster recovery	• Understanding the MetroCluster configuration • Switchover, healing and switchback • Disaster recovery (DR)
Maintain MetroCluster Components	• Guidelines for maintenance in a MetroCluster FC configuration • Hardware replacement or upgrade. Firmware upgrade procedures for FC-to-SAS bridges and FC switches • Hot-adding a disk shelf in a fabric-attached or stretch MetroCluster FC configuration • Hot-removing a disk shelf in a fabric-attached or stretch MetroCluster FC configuration • Replacing hardware at a disaster recovery site in a fabric-attached or stretch MetroCluster FC configuration • Expanding a two-node fabric-attached or stretch MetroCluster FC configuration to a four-node MetroCluster configuration. • Expanding a four-node fabric-attached or stretch MetroCluster FC configuration to an eight-node MetroCluster FC configuration.
Transition from MetroCluster FC to MetroCluster IP MetroCluster Upgrade and Expansion Guide	• Upgrading or refreshing a MetroCluster configuration • Transitioning from a MetroCluster FC configuration to a MetroCluster IP configuration • Expanding a MetroCluster configuration by adding additional nodes
MetroCluster Tiebreaker Software installation and configuration	• Monitoring the MetroCluster configuration with the MetroCluster Tiebreaker software
Active IQ Unified Manager documentation NetApp Documentation: Product Guides and Resources	• Monitoring the MetroCluster configuration and performance
Copy-based transition	• Transitioning data from 7-Mode storage systems to clustered storage systems
ONTAP concepts	• How mirrored aggregates work

Copyright information

Copyright © 2026 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

THIS SOFTWARE IS PROVIDED BY NETAPP “AS IS” AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

NetApp reserves the right to change any products described herein at any time, and without notice. NetApp assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

LIMITED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (b)(3) of the Rights in Technical Data -Noncommercial Items at DFARS 252.227-7013 (FEB 2014) and FAR 52.227-19 (DEC 2007).

Data contained herein pertains to a commercial product and/or commercial service (as defined in FAR 2.101) and is proprietary to NetApp, Inc. All NetApp technical data and computer software provided under this Agreement is commercial in nature and developed solely at private expense. The U.S. Government has a non-exclusive, non-transferrable, nonsublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b) (FEB 2014).

Trademark information

NETAPP, the NETAPP logo, and the marks listed at http://www.netapp.com/TM are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.